EXPERT SYSTEMS APPLICATIONS

for the Electric Power Industry

Volume 1

Joseph A. Naser

Electric Power Research Institute

EXPERT SYSTEMS APPLICATIONS FOR THE ELECTRIC POWER INDUSTRY

EXPERT SYSTEMS APPLICATIONS FOR THE ELECTRIC POWER INDUSTRY

VOLUME 1

Edited by

Joseph A. Naser
Electric Power Research Institute
Palo Alto, California

Sponsored by

ELECTRIC POWER RESEARCH INSTITUTE
Palo Alto, California

HEMISPHERE PUBLISHING CORPORATION
A member of the Taylor & Francis Group

New York   Washington   Philadelphia   London

EXPERT SYSTEMS APPLICATIONS FOR THE ELECTRIC POWER INDUSTRY

1 2 3 4 5 6 7 8 9 0 EBEB 9 8 7 6 5 4 3 2 1 0

Cover design by Renée E. Winfield.
A CIP catalog record for this book is available from the British Library.

Library of Congress Cataloging in Publication Data

Expert systems applications for the electric power industry / edited by Joseph Naser : sponsored by Electric Power Research Institute.
p. cm.
Includes bibliographical references and index.

1. Electric power-plants—Data processing—Congresses. 2. Expert systems (Computer science)—Industrial applications—Congresses. 3. Nuclear power plants—Data processing—Congresses. I. Naser, Joseph. II. Electric Power Research Institute.
TK1191.E974 1990
621.31′0285—dc20 90-4717
CIP

ISBN 1-56032-102-4
EPRI: NP-6957

Contents

TECHNOLOGY, TOOLS, AND METHODS

NUCLEAR POWER PLANT APPLICATIONS

ELECTRICAL SYSTEMS APPLICATIONS

NUCLEAR POWER PLANT APPLICATIONS

ELECTRICAL SYSTEMS APPLICATIONS

VOLUME 2

POSTER PRESENTATIONS

FOSSIL POWER PLANT APPLICATIONS

NUCLEAR PERFORMANCE APPLICATIONS

APPLICATIONS

METHODOLOGIES

TUTORIAL SESSION

Preface

The Nuclear Power, Generation and Storage, and Electrical Systems Divisions of the Electric Power Research Institute (EPRI) sponsored the Conference on Expert System Applications for the Electric Power Industry, which was held in Orlando, Florida, on June 5-8, 1989. The conference was hosted by Florida Power Corporation and Duke Power Company. It was attended by a diverse group of over 300 representatives of electric utilities, equipment manufacturers, engineering consulting organizations, universities, national laboratories, and government agencies. It consisted of a keynote address, 90 papers, 5 tutorial presentations and 3 luncheon presentations by authors from 13 countries. In addition, 25 application systems were demonstrated in the evenings. EPRI has performed and sponsored a substantial effort in advancing the field of expert systems for the electric power industry. Thirty-three papers and 12 demonstrations presented at this conference discussed EPRI-related activities.

Experts from 15 countries were brought together to discuss expert systems applications in the electric power industry. The results of a survey at the end of the conference showed that attendees were impressed with the wide variety of applications that exist or are being developed for the electric power industry. The conference described many expert systems that have already been tested and implemented or are currently in an advanced stage of development. This focus on production grade systems may be contrasted to a meeting just two years ago, when most applications were in the planning or early developmental stages. Thus, this conference marks a major step forward in expert system technology for the electric power industry.

The purpose of this technology transfer conference was to stimulate vigorous efforts to deploy expert system technology by increasing a large and diverse awareness of the number and variety of expert system applications available to the electric power industry. The participants left the conference with a sense of excitement that expert system applications have matured enough to offer immediate and substantial benefits for the electric power industry in a wide variety of domains, including operations, maintenance, and planning. These benefits include increased productivity and efficiency, improved quality, enhanced safety, improved consistency and objectivity, reduced costs, and finally, improved methods for capturing, packaging, and distributing corporate expertise.

Joseph Naser

SESSIONS

KEYNOTE ADDRESS

Expert Systems: A Glimpse into the 1990s

ROBERT S. ENGELMORE
Knowledge Systems Laboratory
Computer Science Department
Stanford University
Stanford, California 94305, USA

I'm very pleased to be given the opportunity to talk about my favorite subject, artificial intelligence, and, in particular, the subfield commonly known as expert systems. Over the next three days you will have the opportunity to hear how expert systems are being used in the electric power industry. Joe Naser has noted that there are 94 papers and two dozen poster presentations in the program. That's a clear indication that the industry is beginning to recognize the value of this technology.

Since you will be hearing so much about what's going on in your domain, I will talk about some applications in other areas, evaluate where we stand today with the technology that's in commercial use, and then tell you about some recent work in our laboratory which is aimed at making these expert systems even better.

Last year, my colleagues, Ed Feigenbaum and Penny Nii, and science writer Pamela McCorduck, published a book called *The Rise of the Expert Company* (Feigenbaum et al. 1988). Written for a non-technical audience, the book is a collection of stories about expert systems which have been developed and put into operation in industry, commerce, and government, with examples from Japan, Europe, and Australia as well as the United States. If these stories are representative of the world at large--a reasonable assumption in my opinion--we are in the midst of an important revolution in the way that organizations are doing their work. They report returns on investment for "small and even medium size expert systems that were in the thousands of percent." One of the big surprises was the almost universal report that these systems were reducing the time to accomplish a task by factors of ten or more. Anytime you gain an order of magnitude in something, you see qualitative changes as well (jet planes are an order of magnitude faster than automobiles, which are an order of magnitude faster than walking). Improved quality of products and/or greater consistency in their manufacture was also evident. Expert systems, as you know, are repositories of the knowledge of experienced specialists. These knowledge bases comprise a sort of corporate memory, ranging from how to troubleshoot a complex device (which the company may no longer manufacture), to how to assess risk in financial operations, to how to optimize the process flow on a shop floor or on a semiconductor fabrication line. Instead of putting this knowledge in bulky user manuals that no one wants to read, the knowledge is preserved in an active medium and made available as it's needed for a particular situation.

Here's a capsule summary of a few stories from the book:

1. Northrop Aircraft in California is using a system called ESP to help process planners plan the manufacture of parts for jet fighters. Today's jet fighters require about 11,000 different types of parts, each of which requires a manufacturing plan, and the parts must be assembled according to an assembly plan--there may be over 20,000 plans in all. With ESP, the process planners report a 12- to 18-fold productivity gain; one person can now do the whole job; and those plans are now generated with greater consistency than ever before.

2. IBM's plant near Burlington, Vermont is using an expert system called LMS to increase the productivity of their microchip production lines. LMS advises operators and managers on the relative priorities of work in the queues, on ways to reroute work if a problem develops at one of the workstations, and sends messages upstream and downstream of the problem, advising the other workstations of schedule changes. It can do some tasks better than humans, such as optimizing the time to shut down the line so as to minimize rework, or to explore alternative line controls to get "the right amount of the right part numbers out every single day." LMS gives managers an overview that they never had before. Although IBM won't release the data, best estimates are that LMS has realized a productivity gain in the tens of millions of dollars per year.

3. American Express uses an expert system called the Authorizer's Assistant at their operations center in Fort Lauderdale. AA not only helps the credit authorizers make their decisions more quickly, but more importantly it helps them make better decisions, decisions which significantly reduce losses to the company by declining bad transactions, and increase revenue by approving good ones. Annual savings here are also in the tens of millions. A number of institutional obstacles at American Express nearly sabotaged the project and I recommend your reading this story to learn some of the many ways an expert systems development project might fail.

4. Here in Orlando, Westinghouse's Diagnostic Center sells a service comprised of a suite of diagnostic expert systems for the major parts of steam turbine generators. Since the rules used in each of these systems come from the best experts in the field, the utilities that purchase this service are getting the very best diagnostic advice available, 24 hours a day. The payoff is increased uptime, 0.9 percent over a recent two-year period. That's about three and a half days per year, and I don't need to tell this audience the cost of a single day's outage. The cost for this service is well below 10 percent of these savings.

5. Canon Research Laboratories in Japan uses an expert system called Optex to assist lens designers. The designer states his goals to Optex, which later works out the details and presents a design. The system can run a complex ray-tracing CAD system and evaluate its designs with respect to the design goals as well as manufacturability. The benefits of Optex are five-fold:

1. It saves time
2. Because it's fast, the space of designs can be explored more fully to find an optimum in performance per unit cost.
3. Patent data can be generated automatically.
4. Programming costs are reduced by reusing and modifying old designs, or subsets of old designs.
5. The designer can explore totally new designs that were previously too costly.

Although cost savings to Canon are substantial--a figure of $700K per year is given in the book--the real payoff is in "working smarter," that is, Optex makes it possible for the lens designers to be truly innovative. When you can generate a design in 15 minutes that used to take three hours to do, you can now test all sorts of ideas that were previously too time consuming or costly to consider.

This is just a small sampling taken from *The Rise of the Expert Company*. There are lots more stories, of course, and in fact, most of them are not in the book. These systems can mean a significant competitive edge for a company, and the authors found (and I've found it true myself) that many organizations will not discuss their expert systems activities publicly, at least not until they're sure they have a significant head start on the competition. We do know, however, that this technology has proven to be useful in a wide variety of human activities. As of mid-1989, we conservatively estimate that there at least 3200 expert systems in actual use (approx. 2000 in the United States, 600 in Japan and 600 in Europe). These system have proven to be useful in all manner of tasks: advisory assistance, configuration, cost estimation, data interpretation, design, diagnostics, emergency procedures planning, financial decisions, insurance underwriting, office procedures, production planning and scheduling, process control, sales, and social services, to name a few.

So, to summarize, expert systems have proven to be a powerful technology that's scoring impressive productivity gains and cost savings, and even allowing some companies to engage in new business areas or to innovate in ways that were previously impractical. But the systems encapsulate only slivers of the knowledge, are only good for doing one thing well, exhibit neither commonsense knowledge of the real world nor any ability to reason from first principles, and generally do a mediocre job of explaining how they know what they know.

One should keep in mind that the commercial systems of today are built upon the research of ten years ago. So, if we want an idea of what the expert systems of the late 1990s will look like, we should pay attention to what's going on in the research labs today. I come from one of those research labs so I'd like to tell you a little bit about our current work there. I make no claims to giving you an overview of the current state of AI research or even knowledge-based systems research in the world today. There's a lot of interesting and relevant work in progress at such places as IBM Research, MIT, CMU, Ohio State, MCC, University of Illinois, and Xerox PARC, among others, but I have neither the time nor the ability to summarize that work here. What I will do is give you a sort of tunnel-vision view into the future and talk about one project.

Under sponsorship from NASA, IBM, and just recently, DARPA, our group, the Heuristic Programming Project at Stanford has been looking at ways to overcome some of these problems I've mentioned, particularly the brittleness of current expert systems and the lack of reusability of their knowledge bases. We were not particularly interested in building an enormous knowledge base that would contain all sorts of commonsense knowledge of the sort that lets us figure out how to get from San Francisco to Orlando if you miss your plane. That's an enormous a task which we'll leave to MCC where Doug Lenat and his colleagues are halfway through a ten-year project, called CYC, to build such an encyclopedic knowledge base, or to the Electronic Dictionary Project in Japan. We decided to focus on scientific and engineering knowledge, where the concepts and relations are less ambiguous, where we feel there's a chance of standardizing the structure and content of the knowledge base, and where we see potential value for the nation's overall productivity within the next decade.

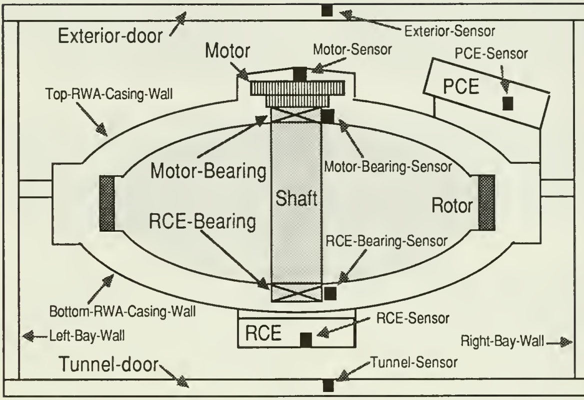
So, where do we start? We started looking at the problems of reusability and brittleness. Could we build a single knowledge base for, say, some electromechanical device from which we could perform more than one task? NASA provided us with an interesting testbed--the Hubble Space Telescope. Since the telescope as a whole is very complex, we focused in on one subsystem called the Pointing Control System, and within that, an interesting device called the Reaction Wheel Assembly (RWA). The HST does not use jets of propellant to turn the telescope, because the propellant might damage the surface of the mirror. Instead, a set of gyroscopic wheels, oriented along different axes, are spun up, and the telescope conserves angular momentum by turning.

The task we set for ourselves was to develop a knowledge base for the RWA that is sufficiently general to allow us to perform at least two different types of tasks. We chose diagnosis and redesign as our initial two tasks. In particular we looked at the problems of diagnosing the cause of overheating indicated by a sensor and at developing a plan for redesigning the RWA to obviate this problem in the future.[1]

One virtue of today's expert systems is that they solve problems efficiently, using, for example, associational rules that directly link symptoms with causes without a long chain of analysis. Having to resort to a general-purpose knowledge base, i.e., to "first principles", on the other hand, would be a tedious way to solve every problem. So we don't want to give up the shallow but very efficient associational rules of today's task-specific expert systems.

---

[1] I am indebted to my co-worker Richard Keller, who is responsible for much of the work reported in the remainder of this paper, and for supplying the figures used here. Readers can find additional detail in [Keller, 1989 #494].

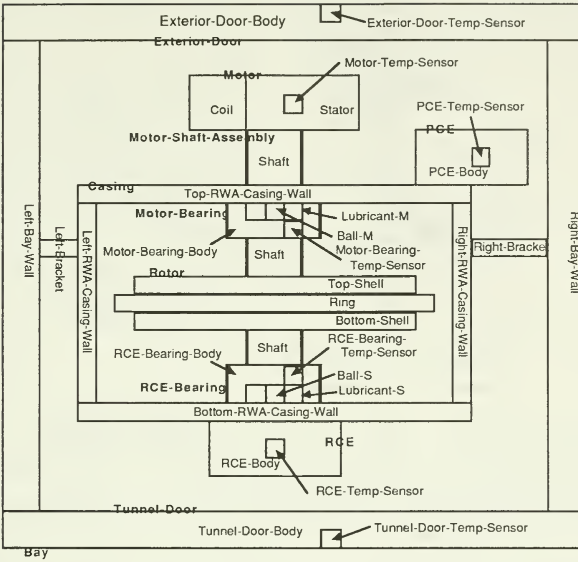


A schematic view of one of the reaction wheel assemblies used to point the Hubble Space Telescope.

Our approach is to develop general-purpose models in a domain, and also to develop knowledge compilation techniques--ways to transform this general knowledge into task-specific rules which can be input to task-specific inference engines.

The model of the RWA has two parts--structural and behavioral.

The structural part is represented in a standard way, using a frame-based, object-oriented knowledge-representation tool (Hyper Class). We represent components, subcomponents, physical connectivity, and spatial relationships. In our initial prototype, we used a two-dimensional boxlike representation which captures the general size and layout of the components, as shown below.



Two-dimensional spatial representation of the RWA.

The behavioral part consists of a set of equations which specify constraints among the parameters which describe the components. The equations may be a mix of quantitative and qualitative relations.

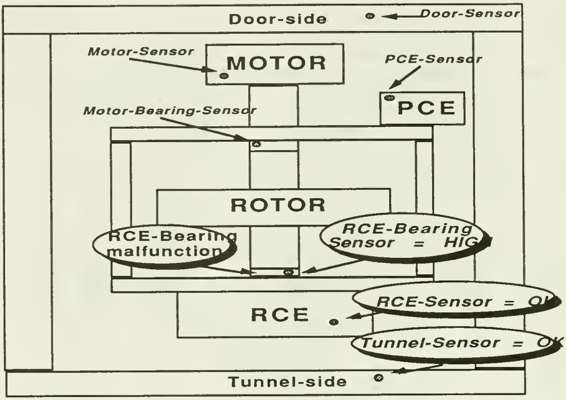
To reiterate, our goal was to demonstrate multiple use of the general knowledge base by compiling the device model into rules for diagnosis and into plans for redesign.

From device models to diagnostic rules

Here's an example of a fault localization rule in a diagnostic system for the RWA:

If the temperature reading of RCE-bearing-sensor-3 is high, and
if the temperature reading of RCE-sensor-34 is OK, and
if the temperature reading of tunnel-sensor-101 is OK,

then RCE-bearing-6 is malfunctioning.



Visualization of the example diagnostic rule

Two things are worth noting about this rule. One, if you consider the structural model as shown in the figure above, you can see that the rule omits sensor readings at other nearby components. These are potential heat sources. Why aren't they considered? The experts who generated this rule considered these other sources to have negligible influence. Today's expert systems would not be able to give you that explanation. When we asked the expert for an explanation, we found that the rule can be justified on the basis of normal processes of heat flow (plus the assumption of correctly functioning sensors). This led us to the development of a model of heat flow within the RWA, which I'll discuss in a moment.

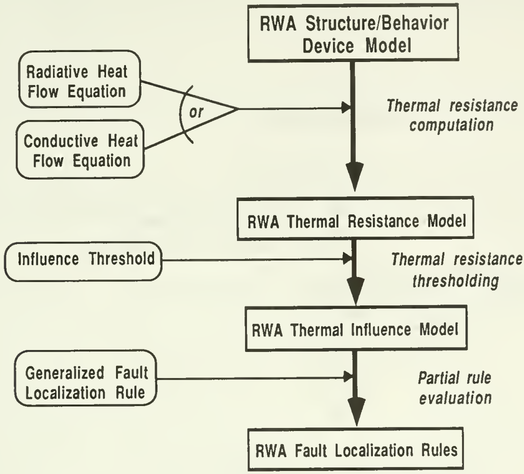
The second thing worth noting is that the rule is a special case of a more general fault isolation rule. Suppose we have a system with a set of n components that are potential sources of problems, and a set of sensors associated with each source. Then we can state the general fault isolation rule as:

If the reading of Sensor(i) is abnormal, and
for all Sources(k), k ~= i, where Source(k) influences Sensor(i):
if the reading of Sensor(k) is normal,

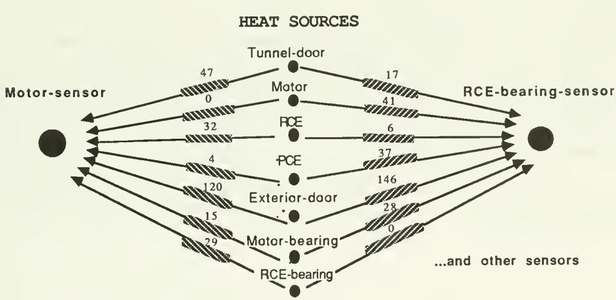
then Source(i) is malfunctioning.

We can get from this general rule to the more specific rule shown on the previous page by using knowledge specific to the RWA device--knowing all the sensors and corresponding sources, and knowing what it means for a sensor value to be abnormal or normal. We also need to know the identity of all heat sources and whether they can "influence" the RCE-bearing sensor.

The overall process of generating a specific diagnostic rule is shown in the figure below. We can derive a thermal influence model from the general-purpose RWA model in two steps. The first step is to produce a simple heat transfer model which uses the concept of thermal resistance. In this model, heat flows along every physical path (by conduction or radiation) between heat sources and heat sensors. The amount of heat reaching a sensor along each path is determined by the thermal resistance of that path, a number that presumably could be derived from a quantitative analysis of heat flow within the RWA structure. Note that this model captures the proper thermal relationships between the components, but loses all spatial relationships.

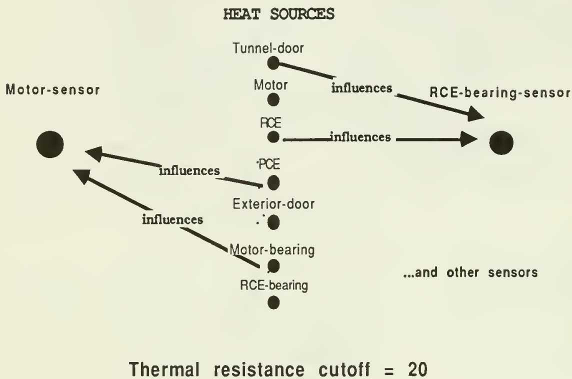


Steps in knowledge compilation for the RWA target diagnostic rule.



Step 1: Thermal resistance model (simplified to show only two of the sensors).

The second step is to define the concept of influence. This can be done very simply by using numerical thresholds. That is, if the thermal resistance between a heat source and a heat sensor is below a certain value, then that source influences that sensor. Note that we lose additional information by taking this step, in that the sensors are no longer "aware" of any components other than those which influence them.



Step 2: Thermal influence model generated by choosing a particular thermal resistance threshold.

Finally, we can produce the target rule we originally wrote down by instantiating the general fault localization rule, using the thermal influence model just derived. Each step in this knowledge compilation process loses information about the device as a whole, but we end up with the efficient, specialized rules that are associated with expert systems. However, we now have a set of models from which the final rule was derived, and we can justify the rule by reinvoking these models. Moreover, we can see how to modify rules automatically if, for example, the structure of the device were changed, thereby changing the thermal resistance values, or if we wanted to examine more subtle thermal influences by raising the thermal resistance threshold.

From device models to redesign plans.

Our second chosen use for the general-purpose RWA knowledge base is for generating redesign plans. To make this more concrete, here is an example of a plan that would be the output of our knowledge compilation process:

If goal is to decrease temperature of RCE-bearing-6,

then (in order)

increase width of RCE-bearing-6
increase thickness of casing -wall-49
increase thermal constant of casing-wall-49
increase width of RCE-body-23
increase thermal constant of RCE-body-23.

Note that this plan is an abstract one. It says what to do, not how to do it, nor does it give any quantitative values (e.g., how much to increase the width of the bearing). However, if we can get this far, there are tools which can use such plans as input and interactively produce more detailed plans.

To derive redesign plans, we use a five step compilation process, which I'll illustrate with the above plan as a target. The first step is to assemble a set of qualitative equations which model the relevant behavior. This behavioral model forms the basis of our redesign plan. We can infer from it what values can be modified and how to modify them to achieve a particular redesign goal. Part of the equation set of interest is shown below.

[BearingTemp6] =[TunnelContrib3] + [RCEContrib4] + [MotorContrib1] + [BearingFriction6]

[MotorSpeed6] = [BallRadius2]+[BearingFriction6]

[MotorSpeed6] = [MotorCurrent3]

[MotorCurrent3] = [RCETemp6]

[BallRadius2] = [BearingWidth7]

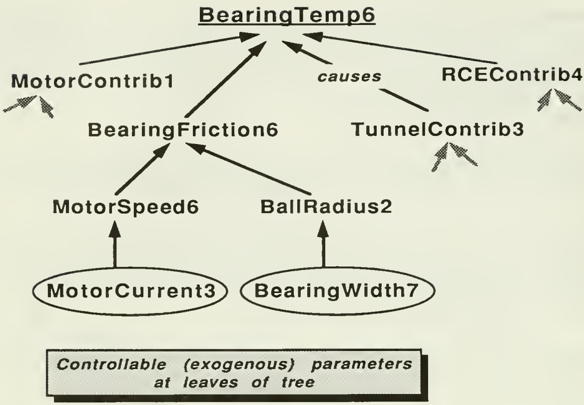
[MotorCurrent3] = [CoilRadius2] + [MotorTemp8]

[DoorTemp2] = [AluminumReflectivity3]

...etc.

Step 1: Equation Set Assembly

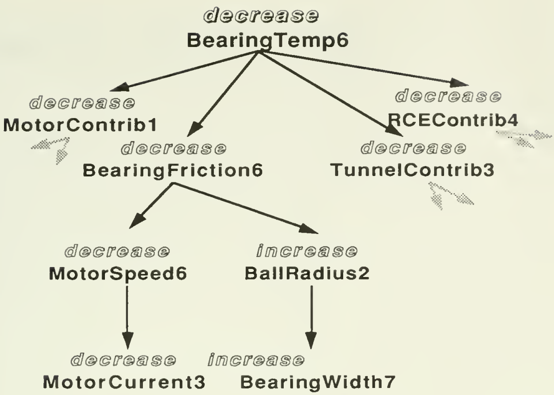
Given these qualitative equations we can use Iwasaki's causal ordering procedure (Iwasaki and Simon 1986) to analyze the causal dependencies. This second step requires specifying which quantities are *exogenous*, i.e. quantities whose values are not determinable from any quantities within the scope of the system under study. These quantities will then appear at the leaves of a dependency graph. Space does not permit an explanation of the causal ordering scheme and the reader is referred to the papers of Simon and Iwasaki for details. The important point to remember is that we can construct a complete causal dependency graph via an iterative process. The figure below shows a portion of the graph, showing the causal dependencies for the quantity of interest in our example.



Step 2: Causal dependency analysis

Note that the causal dependency graph throws away the qualitative relationship between quantities. For example, we can't tell if increasing the radius of BallRadius2 will increase or decrease BearingFriction6 from the graph alone. However, by going back to the qualitative equations, we can change the labels on the arrows from "causes" to "increases" or "decreases". Now we have a redesign goal tree, as shown on the next page.

The fourth step is to prune and order the nodes, and this process usually requires task-specific redesign heuristics. Two types of heuristics are used in our current compiler. One prunes those goals or sub-goals which would violate any given constraints. We may not be allowed to decrease the motor current, for example, because that would reduce the motor torque below a minimum threshold. The second type of heuristic is specific to the thermal model which we introduced when discussing the diagnostic compiler. Thus, if the thermal contribution from the tunnel has a thermal resistance above some threshold, we can prune that branch of the tree. After pruning one can reorder the recommended actions according to increasing thermal resistance.



Step 3: Redesign goal tree generation

The final compilation step is to synthesize the abstract redesign plan. This is a straightforward procedure, in which the root of the tree becomes the antecedent (the condition for applicability of the plan) and the ordered leaves of the tree are the recommended redesign actions. The result is the plan that we wrote down at the beginning of this section.

Conclusion

The work at our laboratory is still in an early stage of progress and I don't want to make any strong claims for its generality. However, I think it's in the mainstream of AI research going on today all over the country, research which will give us reasoning systems that are not only knowledgeable, but robust, that can employ that knowledge in multiple tasks, and that can justify their conclusions on the basis of models of their domain at different levels of abstraction.

References

Feigenbaum, E. , P. McCorduck and H. P. Nii. (1988). The Rise of the Expert Company. New York, Times Books.

Iwasaki, Y. and H. A. Simon. (1986). "Causality in Device Behavior." Artificial Intelligence. **29**: 3-32.

GENERAL OVERVIEW

EPRI's Nuclear Power Division Expert System Activities for the Electric Power Industry

JOSEPH A. NASER
Electric Power Research Institute
Nuclear Power Division
3412 Hillview Avenue
Palo Alto, California 94304, USA

ABSTRACT

Expert system technology has demonstrated its capabilities and benefits in a broad range of applications and domains. Three major goals of high technology applications for nuclear power plants have been identified by an advisory group of utility personnel. These goals are to enhance power production, to increase productivity and to reduce safety challenges to the plant. The ability of expert systems to enhance productivity, to aid in decision-making and to capture and distribute corporate expertise make them an important technological tool for the electric power industry for achieving these goals.

Two parallel efforts are being performed by the Nuclear Power Division of the Electric Power Research Institute (EPRI) to help the electric power industry take advantage of this expert system technology. The first effort is the development of expert system building tools which are tailored to electric power industry applications. The second effort is the development of expert system applications. The purpose of this paper is to describe some of the tool and application development work which is being performed by the Nuclear Power Division for the electric power industry. This work includes prototypes developed to demonstrate feasibility, production systems under development and systems which have been implemented. This paper will also describe some of the other efforts such as the development of the material for a knowledge acquisition workshop, the development of expert system verification and validation methodologies and the use of expert systems themselves for technology transfer of EPRI research results.

INTRODUCTION

Research in the field of Artificial Intelligence (AI) has been going on since the mid 1950's. This research includes robotics, modeling the human cognitive processes, vision, speech, natural language processing, theorem proving, automatic programming and expert systems. The modeling of human cognitive processes for solving significant problems by trying to duplicate the behavior of the human brain was not initially very successful due to the lack of sufficient computational power. As an alternative approach for solving significant problems, the concept of an expert system was developed. Edward Feigenbaum, a pioneer in the field of expert systems, developed the key idea that knowledge is power and that the more knowledge, the more powerful. Expert systems are an embodiment of this concept. They contain knowledge of the domain, usually in a symbolic representation, and reason about that knowledge symbolically.

The first expert systems emerged in the late 1970s. Researchers at Stanford University developed MYCIN, the first interactive consultative expert system, for bacterial infectious disease diagnosis and therapy, and DENDRAL, the first expert system, for computing structural descriptions of complex organic chemicals. Digital Equipment Corporation (DEC) developed R1 (later renamed XCON) for determining specifications and configurations for DEC's computer hardware. Schlumberger Ltd. developed the Dipmeter advisor for analyzing geological formation encountered in oil well drilling. These systems led to an explosion of expert systems in the 1980's. As of 1989 it is estimated that there are over three thousand expert systems, of which about two thirds are in the United States. These applications range from very simple to very complex ones and include all sectors of industry. This expert system explosion grew out of the perceived and realized benefits of expert systems. These benefits include increased productivity, improved quality, improved consistency, reduced costs and captured corporate expertise. The ability of expert systems to capture knowledge and distribute it has led to substantial increases in revenue and cost savings. These benefits are described in The Rise of Expert Company [1] for such companies as IBM, DuPont, DEC, American Express, Westinghouse, FMC, Canon and others.

The obvious capabilities and benefits of expert systems and their potential to help the nuclear power industry, and the electric power industry in general, was realized by the EPRI Nuclear Power Division in late 1983. At that time the Control and Diagnostics Program in the Nuclear Power Division of EPRI initiated two parallel paths for developing expert system technology to respond to electric utility needs. The first is the development of expert system building tools which emphasizes electric utility applications. The second is the development of expert system applications for the electric power industry. These applications build on the electric utilities' knowledge bases. Each effort provides useful feedback for the other. The application developments help identify the capabilities needed for building expert systems. In addition, the application developments help test the expert system building tools and identify their limitations. The expert system building tools help identify the types of applications which can be successfully developed using a tool. The use of a tool increases the efficiency of the development efforts and helps reduce the costs of development. It also helps to identify and explore the possible knowledge structures and reasoning strategies for the application domain.

Expert system (or knowledge-based) technology has a number of unique capabilities which makes it an important computer resource for the electric power industry. These include programming flexibility, which allows rapid development and modification; inference capabilities, which allow reasoning to be performed in a non-procedural manner over facts and heuristics; explanation facility, which allows the user to ask how a result was obtained; and knowledge structured according to human models, which allows easier understanding and verification of the internals of the expert system. Expert systems can be used as an assistant, a colleague or an expert consultant for the user. They create a benefit to the electric power industry by capturing, refining, packaging and distributing expertise; preserving the utility's knowledge; solving problems more quickly and efficiently; solving problems more objectively and consistently; solving problems which require the knowledge and expertise of several domains; solving problems where the required scope of knowledge exceeds that of any single person; and solving problems whose complexity exceeds human ability. Each of these capabilities of expert systems can help achieve the goals of enhancing power production, increasing productivity and reducing safety challenges to the plant which were set by the EPRI Nuclear Power Division's Control and Diagnostics Utility Subcommittee.

Another area of expert system technology work being performed by the Nuclear Power Division is technology transfer. This includes the development of workshops to transfer expert system technology to the electric utilities and the use of expert systems as a means to transfer EPRI research results to the electric utilities. Research is also being performed on the development of verification and validation methodologies for expert systems to enhance their acceptance by users and regulators.

Expert system technology represents another computer tool which is available for solving problems. In spite of the somewhat imposing name, expert systems are really just intellectual assistants and intellectual power tools for the users. They more often play the roles of colleague, assistant and servant than expert. After understanding that expert systems are very powerful tools, which should be used when needed, it is appropriate to consider areas where expert systems technology might be applied usefully in the electric power industry. These areas include diagnosis, monitoring, interpretation, instruction, planning and prediction. In order to capitalize on the benefits, which can be achieved by expert systems in these areas, the Nuclear Power Division has been developing the expert system building tools and applications described below.

EXPERT SYSTEM BUILDING TOOL DEVELOPMENT

The EPRI program to develop expert system building tools includes five development projects for development of PLEXSYS, SMART, ProSys, IRTMC and TRESCL. These tools cover a wide range of expert system capabilities as will be described below.

The objective of the PLEXSYS (PLant EXpert SYStem) [2,3] project is to develop a specialized expert system software tool for electric power industry applications which facilitates expert systems development by electric utilities and their suppliers. This software tool will be especially suited for nuclear power plant expert systems involving plant design, engineering and maintenance activities. It is equally applicable to other types of power and process plants.

This development effort is based on extensions to the commercial artificial intelligence toolkit Knowledge Engineering Environment[4] (KEE). Since expert system tools are a rapidly developing technology, the adaptation of commercial software enables the enhancements of this project to "float" on the technological improvements fostered by other segments of the artificial intelligence research and development community.

PLEXSYS has been developed for expert systems for modeling complex physical systems such as electric power plants. The central facility in PLEXSYS is a model editor which enables users to build or represent their plant in a schematic format similar to computer-aided design (CAD) systems. For example, this allows the user to work with the piping and instrumentation diagram (P&ID) formats with which he is familiar. However, in addition to the schematics are data or "knowledge" base structures and methods which automate reasoning and problem solving tasks involving complex systems. An example of this are the facilities for performing various types of network analyses. PLEXSYS is complete and it has been formally released. An effort is also underway to automate the building of the PLEXSYS knowledge base directly from a CAD data base.

The "Small Artificial Reasoning Toolkit (SMART)"[5] development provides a compact, personal computer-based expert system development toolkit that electric utilities can use to develop a variety of small-scale expert systems applications. SMART was built for standard personal computer systems without requiring special memory or accessory devices. An overlay LISP symbolic programming environment with sufficient built-in, top-level capabilities exists enabling users to construct expert systems without requiring a priori programming experience. SMART was developed to provide knowledge representation, reasoning and interfaces to LISP which allow advanced users to construct sophisticated expert system applications.

SMART supports object-oriented, frame-based knowledge representation with inheritance properties, forward and backward chaining inference methods, embedded methods, query functions, explanation capabilities, demons, interactive menu constructs, and assorted utilities for customizing and extending SMART for specific applications. SMART is complete and has been formally released.

ProSys[6] is a model-based diagnostic expert system environment on a 386 personal computer which is an enhanced and more generic implementation of the National Aeronautics and Space Administration's (NASA) KATE [7] (Knowledge-Based Autonomous Test Engineer) environment. The objective of the ProSys development is to provide a tool which allows the representation of complex physical systems through structural and functional information.

ProSys, as does KATE, inherently knows how to perform the capabilities of system monitoring, signal validation, fault location and diagnosis, automatic control and automatic reconfiguration. It creates a knowledge base of the physical system model in terms of structure and function and uses this knowledge to draw inferences about the current state of the system. ProSys is capable of predicting the expected sensor values from the system state and operator actions. When the measured sensor values are different than the expected ones, the system determines and diagnoses the failed component or sensor. The first level of ProSys development is complete and is being released for use by the electric utilities.

The Intelligent Real-Time Monitoring and Control Architecture (IRTMC) project is developing a generic architecture which could be used as a platform for various real-time expert system applications. The objective is to develop a system which would acquire data automatically, synthesize data into a dynamic model of the system's functioning, and dynamically plan effective programs for appropriate action. It would integrate quick, reactive responses to urgent events with carefully planned courses of action for managing evolving situations. Acting in the role of an intelligent consultant, it would explain its observations, reasoning, conclusions and recommendations. In appropriate circumstances, it could perform closed-loop control.

IRTMC will consist of a collection of capabilities which are built on the BB1 blackboard control architecture [(8)]. The BB1 blackboard architecture provides mechanisms for knowledge representation, reasoning and strategic control. Currently a prototype system for medical intensive-care monitoring is being developed. This project will take the architecture developed for medical applications and develop a generic architecture which is useful in the domain ofpower plants. The generic reasoning capabilities currently include data filtering, data classification, associative diagnosis, model-based diagnosis and reactive response. This work is just beginning.

The objective of the TRESCL [(9)] (Translate Expert System to C Language) is to develop the capability to translate LISP-based expert systems into a high performance C language implementation. This effort is being performed by using SMART as a model for prototyping generalized capabilities. Using a structural approach, C language emulations of the principal SMART functions are being developed. These emulations make maximum use of C language programming constructs and will pre-link rules and other objects for topological search of sematic networks in lieu of rule chaining operations. TRESCL accepts knowledge bases developed with SMART. This tool is at the research-grade level.

EXPERT SYSTEM APPLICATIONS DEVELOPMENT

A number of expert system applications for the electric power industry are being currently developed by the Nuclear Power Division. These are in varying stages of prototype or production system development with some of them implemented and being tested. These applications can be put into three basic categories of expert systems. These categories are Classification, Planning and Diagnosis. The first seven applications to be described fit into the category of classification expert systems.

The first classification expert system is the "Emergency Operating Procedures Tracking System"[10]. The objective of this project is to develop a computerized system to help operators select and apply operating procedures during plant emergencies. This project will provide the capability to interpret and compile emergency operating procedure logic into a compact, fast-running software module that interfaces to and is co-resident with the nuclear power plant's Safety Parameter Display System (SPDS). It utilizes the same data base as does the SPDS. A custom-made inference engine and knowledge representation scheme was developed in C for the emergency operating procedures tracking system. This was done to ensure very high speed and efficient memory utilization by the system. For some applications this approach may be a necessary or desirable strategy instead of using an off-the-shelf expert system shell. The emergency operating procedures tracking system allows multiple user access (e.g., from the control room and the technical support center) and provides real-time notification of emergency procedure steps, on-line explanations for these messages, priority filtering and data quality checking.

The emergency operating procedures tracking system has been fully developed for Boiling Water Reactor (BWR) emergency operating procedures. Initially based on the Boiling Water Reactor Owner's Group emergency procedures guidelines (EPGs), the system has been applied specifically to Taiwan Power Company's Kuo Sheng plant's emergency operating procedures. This system has been implemented as an add-on module to the SPDS developed by General Electric Company for the Kuo Sheng plant. The emergency operating procedures tracking system has been interfaced to the Kuo Sheng full-scale plant simulator for site acceptance testing and performance evaluation by plant operations as a prelude to actual plant installation. Initial testing has indicated that the emergency operating procedures tracking system helps the operators respond in a time indicative of skill-based response instead of knowledge-based response which is achieved without the system.

The second classification expert system application is the "Reactor Emergency Action Level Monitor" (REALM)[11] system. The objective of this project is to develop an expert system for assessing the nuclear plant overall safety situation as an aid to site emergency coordinators. This system interprets the decision logic associated with emergency action levels (EALs) in site emergency response plans.

This expert system captures the expertise and knowledge used by plant technical support personnel as input to the decision logic and rationale embedded in the expert system. This multi-disciplinary approach for assessing the plant condition considers radioactivity release, fission product barriers, critical safety functions, anticipated accidents and safety systems in order to provide reliable emergency action level classifications and supporting rationale over a broad spectrum of plant events.

A full-scale prototype expert system has been developed, using Consolidated Edison's Indian Point Unit 2 as a plant model. The REALM system is presently implemented on a compact workstation using the KEYSTONE[12] artificial intelligence software toolkit. REALM can also be used in a stand-alone configuration for emergency drill scenario development and training applications. The user can test his analysis and decision skills against the expert system with embedded facilities to record and compare the human and machine responses to various emergency scenarios. REALM has been tested off-line at Indian Point Unit 2 during several emergency drill exercises with very favorable results. It is currently being implemented as both an on-line and off-line system at Indian Point Unit 2 and as an off-line training and scenario development tool at Public Service Electric and Gas Company's Salem plant.

The third classification expert system is a "Low Level Waste Advisor". The objective of this project is to develop the specification for and to evaluate the feasibility of an expert system which would be a decision aid for low level waste operations.

Extensive documentation has been developed on low level waste management at nuclear power plants. Since the knowledge which would support any one decision is most likely to be scattered throughout this extensive documentation, this project would develop a system which would aid the rad waste decision-maker by putting all of this knowledge into a single-point control logic system. This system would provide distinct cost, planning, training and regulatory compliance benefits. The development of the specifications is just being initiated.

The fourth classification expert system is LIFEX [13] which provides knowledge-based guidance for determination of potential degradation mechanisms as part of nuclear power plant component life estimation. This system was developed as part of the EPRIGEMS technology transfer program at EPRI. EPRIGEMS has defined a framework and "look-and-feel" on a personal computer which allows expert system technology to be used to transfer results of EPRI research projects to the electric utilities.

LIFEX identifies potentially active mechanisms of degradation over the course of plant life based on the responses to a series of questions. This represents the first step in the evaluation of the remaining life of light water reactor components. LIFEX deals with more than twenty mechanisms that have the potential to influence the performance of LWR structural material. It also includes guidelines which provide utility engineers with the information to assess the potential degradation of plant components. LIFEX is complete and available for use.

The fifth classification expert system is the Safety Review Advisor. The objective of this effort is to help perform safety reviews and 10CFR50.59 reviews for both design and procedure changes. The major effort will be to develop generic rules and to provide guidelines to help electric utilities develop their own plant-specific safety review advisor system.

The requirements for the safety review advisor were identified by an electric utility working group. This system will behave as a smart guide through the review process by using the user's responses to recommend the most relevant topics for further questioning and evaluation. The system will have several options for access to necessary data sources such as the Final Safety Analysis Report and Technical Specifications. This work is just beginning.

The sixth classification system "A Utility's Activities and Research Information System" is designed to look at electric utility activities and available research information to identify potential activities where artificial intelligence techniques may be benefically applied to the operation of nuclear power plants. A methodology will be developed and implemented for identifying and evaluating those activities which could be benefically enhanced by artificial intelligence techniques. The project is currently working on identifying the appropriate attributes of nuclear power plant activities which will help determine the applicability of artificial intelligence techniques.

The last of the classification expert systems is a personal computer-based "Snubber Reduction/ Piping Design Improvement" expert system. This system will guide electric utilities in evaluating the cost-effectiveness of snubber reduction/piping design improvement and in implementing such an effort.

This system will respond to user's questions to give advice on snubber reductions. This advice will be based on the stored knowledge base and supplementary interactive queries. The system will supply information about required analyses, criteria to be met, licensing issues to be addressed and other considerations to be included to achieve maximum snubber reduction. The cost-effectiveness can then be calculated, and procedures to implement snubber reduction/piping design improvement can be defined. This work is just beginning.

There are five expert system applications to be described in the category of planning expert systems. The first of these is a "Refueling Insert Shuffle Planner".[(14)] The objective of this project is to develop the capability to determine an efficient refueling crane movement pattern for the fuel insert shuffle of a Pressurized Water Reactor (PWR) when this shuffle is performed entirely in the spent fuel pool.

Using Virginia Power Company's Surry Units 1 and 2 as a test bed plant model, a knowledge-based system, using the commercial artificial intelligence software KEE, was developed as a full-scale prototype. The technique for developing the crane movement pattern is independent of reactor and spent fuel pool geometries. It is based on building up chains of moves which are independent of each other. Only the graphical user interfaces are site-specific.

The approach used in the refueling insert shuffle planner does not find an optimal solution, since an optimization is believed to be too difficult and time-consuming. Instead, heuristics are used which will find a number of very good solutions. Then the user can select the best of these solutions. Rules are used to allow electric utilities to easily incorporate their specific constraints on the system. This prototype system has been completed and tested.

The second planning expert system is a "Planning System for Core Shuffles". The objective of this system, based on the success of the Refueling Insert Shuffle Planner described above, is to extend the crane movement planning capability into a production system. The core shuffle planning system will be applicable for PWRs and BWRs. It will handle in-core shuffles for PWRs and BWRs and total core off-load spent fuel pool shuffles for PWRs.

This system will allow for interactive modifications of the shuffle plan as well as the automatic generation of the plan. It also has the ability to graphically walk-through the shuffle plan for easy verification. The system is being made as generic as possible to allow easy modification for plant-specific configurations. This development effort has completed knowledge acquisition and development of the man-machine interfaces. The shuffle strategies are now being implemented.

The third planning expert system application is "A Fuel Shuffling Expert System"[15]. The objective of this effort was is to investigate the potential of artificial intelligence techniques in the nuclear power industry by developing a prototype system for efficiently determining fuel assembly configurations to support PWR reload design.

Using rapid prototyping techniques, the approach was to develop an expert system for interactively analyzing fuel assembly burn-up characteristics and for shuffling assemblies to develop case input to the BETCY/PDQ-7 mainframe core physics analysis codes. This system implements methods for automating input preparation, for associating job control language (JCL) files for downloading and running BETCY/PDQ-7 on a remote mainframe, and for uploading mainframe results for further analysis using the fuel shuffling expert system. Simple heuristics and constraint checking rules were developed to demonstrate expert system capabilities.

An initial prototype was developed and demonstrated using the commercial software toolkit KEE. The prototype did not include a full complement of heuristics for automatically generating new core maps, but did establish a conceptual design to demonstrate feasibility of an expert system core reload design workstation. No additional work is planned for this system.

The fourth planning expert system is an "Equipment Tag-Out System". The objective of this project is to develop the expert system capability to automatically create and plan equipment tagouts as an integral part of an electric utility's computer-based work authorization information system (WAIS) for a nuclear power plant.

This project used the PLEXSYS artificial intelligence toolkit described above to build a plant system model for a prototyping application for maintenance planning and equipment tagouts. The residual heat removal (RHR) system at Pacific Gas and Electric's Diablo Canyon plant was the focus for this work. The PLEXSYS model editor was used to build a component model. System functional states were related to the components and rules were developed to represent the Technical Specification's Limiting Conditions for Operation relevant to the RHR system. This protoytpe system has been completed and successfully demonstrated.

The last planning expert system is the "Component Life-Cycle Advisor". This personal computer-based system is to provide guidance, methods, good practices and tutorials for management of component life-cycle costs. The first component selected for this application is the feedwater heater.

This expert system will permit electric utility personnel to benefit from the vast amount of information which has been gathered and documented on the operation and performance of feedwater heaters. It will also produce a generic life cycle advisor which can have the knowledge of any plant component put into it. The system will aid the electric utility management, engineers, and other planning personnel in minimizing life cycle costs. This effort is expected to begin soon.

The next nine applications to be described fit into the category of diagnostic expert systems. The first of these diagnostic systems is a prototype which was developed to transfer expert system technology from the National Aeronautics and Space Administration (NASA) to the electric power industry. This project transferred NASA expert system technology, which is embodied in the Knowledge-Based Autonomous Test Engineer (KATE)[(7)] expert system environment, by developing a comparable expert system environment ProSys [(6)] and a prototype application for a physical system on a nuclear power plant.

The first step in this technology transfer effort was to evaluate a number of physical systems in a nuclear power plant which could benefit from this technology. EPRI worked with ten electric utilities to identify an important application area. The area selected was alarm processing and diagnosis. A prototype system for the reactor coolant pump seal injection system was developed to demonstrate feasibility of the methodology. For nuclear power plant applications the automatic control and reconfiguration will be replaced by advice to the operator on control and reconfiguration. This work has been completed.

The second diagnostic expert system is the "Alarm Processing and Diagnostics System". The objective of this project is to develop an advice system to help plant operators by prioritizing alarms and emphasizing the most significant ones.

The system will use model-based reasoning as well as rule-based heuristics to obtain high confidence alarm processing and diagnostics from real-time plant data and alarm status. The power plant operator's alarm procedures will be used to help guide the system. This expert system will not change the alarm panel behavior in the power plant. It will be an auxillary tool for use by the plant operations staff. A large-scale system is being developed for Pacific Gas and Electric Company's Diablo Canyon plant. This project has completed the knowledge acquisition phase and is now in the implementation phase.

The third diagnostics expert system is the "Emergency Diesel Generator Diagnostics System". The objectives of this project are to increase the availability and reliability of diesel generators, decrease plant shutdown time caused by diesel generators and to reduce the probability of station blackout.

This project is developing an on-line diagnostic system which will determine predictive maintenance needs by anticipating problems. It will also perform the more traditional fault diagnosis as needed. The system is being developed for Duke Power Company's McGuire plant. The knowledge base for the system is being put together from experience over a wide range of diesel generator types to make it as generic as possible. The project has completed the knowledge acquisition phase and is in the initial development phase. The associated on-line monitoring system has been designed.

The fourth diagnostic expert system is "A Plant Thermal Performance Advisor". The objective of this project is to develop a personal computer-based nuclear power plant thermal performance diagnostics expert system. It will also provide guidance to the electric power industry for plant-specific configuration conversion and for modifications and enhancements to its thermal performance knowledge base.

This project will develop a thermal performance advisor knowledge base from previously documented EPRI work(16). This advisor will assist plant engineers and operators to diagnose heat source related problems based on the user's response to a series of questions by the system. It will suggest additional testing or inspection procedures and provide guidance on corrective measures. This project has demonstrated the first level prototype and is currently developing the production system.

The fifth diagnostic expert system is the "Rapid Repair Advisor". The objectives of this project are to develop field grade expert systems for diagnosis of critical plant equipment and to improve plant capacity.

This project will develop a framework for power plant diagnostic applications. The objective is to have a portable system which can be used by the maintenance staff to aid in equipment diagnostics. The framework is being developed to allow the maintenance person to load into a portable computer the appropriate application software for the equipment being diagnosed. The first application to be developed in this framework is a motor-operated valve diagnostic system. Pacific Gas and Electric Company's Diablo Canyon and Pennsylvania Power and Light Company's Susquehanna plants are being used to develop this capability. This project is in the knowledge acquisition phase.

The sixth diagnostic expert system is a "BWR Transient Diagnostic System"(17). The objectives of this project are to demonstrate the feasibility of a diagnostics system to determine the type and cause of a BWR transient and to demonstrate the feasibility of using a transient analysis computer code as a knowledge source for a diagnostic system.

This project used the RETRAN thermal-hydraulic analysis code to develop the plant transient knowledge base. The system uses transient plant data and alarm status as an input to determine the type of transient which is occurring. When needed and possible, information that is not directly measurable, will be deduced from other observables. A separate rules construction was interfaced with the transient diagnostic system to provide a causal simulation of BWR transients. A prototype, which successfully diagnoses thirteen different BWR transients, was developed to demonstrate feasibility.

The seventh diagnostic expert system is a "BWR Shutdown Analyzer"(18). The objective of this project is to investigate the potential of artificial intelligence techniques in the nuclear power industry by developing a prototype expert system for analyzing BWR shutdowns.

Using Tennessee Valley Authority's Browns Ferry Unit 1 as a representative plant model, a knowledge-based system using a commercial artificial intelligence software tool (KEE) was developed as a rapid prototype. Rules were provided to analyze reactor trip conditions and determine whether the occurance was either an anticipated transient without scram, a normal shutdown, or an abnormal shutdown. A separate rules construction was interfaced with the shutdown expert system to provide a causal simulation of BWR shutdown systems capable of representing various combinations of malfunctions. The prototype was completed and established feasibility for prospective production systems.

The eighth diagnostic expert system is a "Secondary Side Transport and Retention of Radioactive Species (STARRS) Analysis Tool". This is a diagnostic system which is built in the EPRIGEMS technology transfer framework. It is developed to help plant engineers and operators diagnose the activity transport and retention mechanisms following a steam generator tube rupture design basis or beyond design basis event. The system is currently being pre-release tested by electric utility personnel.

The last diagnostic expert system, and last expert system application to be described here, is CHEXPERT. This system is being developed to assist users in the evaluation of thinning of pipe walls due to corrosion from flowing water. It is also built in the EPRIGEMS framework.

CHEXPERT considers single- and two-phase erosion-corrosion, cavitation, flashing, microbial corrosion and intergranular stress corrosion cracking. It incorporated training, diagnosis and prediction of in-service degradation in piping systems. The diagnostic feature, based on the information supplied by the user, will help identify the probable cause for a given problem and recommend a solution. This effort is nearing completion.

EXPERT SYSTEM RELATED PROJECTS

In the Nuclear Power Division some additional projects related to expert systems are being carried out. They include development of expert system verification and validation methodologies, knowledge engineering techniques, training and design.

Verification and validation has been used extensively in the nuclear power industry to ensure the quality of the product. Examples include on-line systems such as the SPDS and analysis tools such as RETRAN. In some application areas where expert systems offer considerable benefits, an obstacle to their acceptance by both users and regulators is the lack of verification and validation methodologies. The Nuclear Power Division has initiated research into the development of verification and validation techniques for expert systems.

Considerable work has been done developing verification and validation techniques for conventional software systems. This previous work is being taken advantage of and, where applicable, being adapted or modified for expert systems. Additional verification and validation techniques are being explored to handle the unique characteristics of an expert system's knowledge base and the iterative nature of the expert system development process. These unique characteristics include the need to be able to certify the expertise which is being put into the expert system. A method for developing validation scenarios is also being explored. The first steps of the research to develop detailed verification and validation methodologies for expert systems are documented in two EPRI reports.[19,20]

Another area of importance is knowledge engineering, that is, the acquisition of knowledge and its representation in the expert system. This step is frequently considered to be the bottleneck of expert systems development, as expert systems are only as powerful as the knowledge they contain. In most cases this knowledge exists with electric utility personnel who are not expert system developers. Therefore, it is important to develop techniques which will help acquire this knowledge in the electric utility environment. Techniques for knowledge acquisition and representation have been gathered and documented in an EPRI report.[21] In addition, two workshops on these topics have been given to electric utility personnel. An area where expert systems offer considerable promise is the role of an intelligent tutor that is always available when required. An intelligent tutor could also allow the user to proceed at whatever pace is comfortable and backtrack as desired.

The potential of and guidance on the use of expert systems as intelligent tutors has been explored using the REALM expert system as a case study [22]. This effort developed detailed descriptions of expert training system models such as basic domain, trainer and trainee models. Guidelines for developing expert training systems were assembled.

The last project to be discussed in this paper is one to explore the interfaces between computer-aided engineering (CAE) and expert systems. The objective is to combine the graphics and data base capability of modern CAE systems with expert reasoning to capture the expertise of the original system designer, to extend available design expertise using expert systems technology to supplement less skilled designed personnel, to preserve design expertise, and to automate routine design tasks by providing embedded capabilities for intelligent reasoning. So far the project has completed a literature review and a survey of the industry working in this area. A prototype of a reactor design system is being developed.

CURRENT EXPERT SYSTEM TECHNOLOGY LIMITATIONS

As illustrated by the wide variety of expert systems described above, it is obvious that expert system technology has matured enough to be very beneficial to the electric power industry. However, there are still a number of limitations to expert system technology which prevent certain types of applications from being developed. Some of the areas which are still in the artificial intelligence research area are:

- large-scale real-time process control systems;
- very large-scale complex planning systems;
- multiple cooperating intelligent agents;
- large-scale real-time simulation systems;
- large-scale real-time predictive systems;
- pattern recognition systems including speech and vision;
- rigorous and practical handling of uncertainty;
- nonmonotonic reasoning and truth maintenance systems;
- learning and adaptive systems; and
- self-knowledge about limitations of the expert system's capabilities.

As the research efforts bear fruit in these areas, the range of possible expert system applications in the electric power industry will grow. For example, on-line predictive maintenance systems will be more useful and powerful with the inclusion of robust techniques for pattern recognition. These systems will be able to look at the raw data from sensors and determine patterns which would be used by the diagnostics portion of the system.

Considerable efforts are being put into these research areas by the artificial intelligence community. The work on IRTMC with Stanford University is an example of this for one area. As progress is made in these areas, the technology will be incorporated into the electric power industry for additional and more powerful applications development. In the meantime, the current technology is already powerful enough for substantially beneficial applications in the electric power industry.

CONCLUSIONS

This paper has described a number of research projects which are being performed by the Nuclear Power Division of EPRI in both the areas of expert system building tool development and expert system application development. These two parallel development paths have been very beneficial to each other by supplying feedback to each other. The wide variety of expert system applications described here demonstrates a portion of the wide-ranging capabilities of expert systems to assist the electric power industry. Other divisions of EPRI and other organizations are also developing expert systems for the electric power industry. From the work that has already been performed with expert systems in a variety of application areas for the electric power industry, it is obvious that expert system technology is capable of helping electric utilities satisfy their goals of enhancing power production, increasing productivity and reducing safety challenges.

Artificial intelligence in the form of expert systems, as demonstrated by the developments described above, has been established as a credible technological tool for the electric power industry. Expert systems are a method for preserving an electric utility's knowledge base, which is an important part of its corporate assets. Expert systems are useful in a wide, diversified set of applications. Artificial intelligence is a powerful and logical extension of computer power for plant operation, plant engineering and emergency management. A number of expert systems are being developed either as demonstration prototypes or as production systems, and the first applications have only been recently completed and are being used by the electric power industry.

Expert systems have the potential to be useful in a wide range of application areas. Expert system technology is currently not capable of supporting all of the application areas that could benefit from it. Some of the areas, which hold a great deal of promise, are large-scale real-time process control, large-scale cooperating systems, large-scale simulation and predictive systems, and learning systems. A commitment to extensive research and application development in these and other areas are needed to help the technology mature and realize its full potential. In addition, work must be done to develop industrial grade applications and delivery vehicles for these expert systems to be useful in the electric power industry environment. In order to enhance both user and regulatory acceptance, verification and validation methodologies for expert systems must be developed. Some initial efforts have been made in this area with additional work being initiated.

An additional challenge is to transfer expert system technology and an understanding of its potential to the electric power industry. It is not adequate to develop applications and give them to the electric power industry to use as a completed system. First of all, most expert systems will need to be tailored to each electric utility's needs. Second, the nature of these systems is that knowledge should be added to the expert system to enhance its capabilities as the electric utility learns more about the physical system. Also because expert systems hold so much potential in so many areas, the electric utilities will need to develop their own expert systems. This is why the Nuclear Power Division of EPRI is putting extensive efforts into developing a methodology for identifying expert system enhanceable activities into tool development and into technology transfer activities as well as into applications development.

Expert systems have already proven their value in a broad range of domains in other industries. For many applications the quantified benefits from these expert systems is enormous and is measured in terms of millions of dollars in either savings or increased revenue. [1] These systems have been shown to amplify people's capabilities by a factor of ten or more. The Nuclear Power Division is striving to make these types of benefits available to the electric power industry.

ACKNOWLEDGMENTS

The work described in this paper is the work of a number of my colleagues at EPRI as well as my own work. I would like to acknowledge Bill Sun, David Cain, Robert Colley, Norris Hirota, Glen Snyder, Floyd Gelhaus, H. T. Tang, Jeff Byron, Mel Lapides, Pal Kalra and Bindi Chexal for their work on expert system in the Nuclear Power Division.

REFERENCES

1. E. A. Feigenbaum, P. McCorduck and H. P. Nii. The Rise of the Expert Company, New York: Times Books, 1988.

2. W. S. Faught. "Functional Specifications for AI Software Tools for Electric Power Applications." EPRI NP-4141, August 1985.

3. S. Hashemi et al. PLEXSYS Plant Expert System Development Environment, EPRI Report to be published.

4. "Intellicorp KEE Software Development System User's Manual." KEE Version 3.0, Intellicorp, July 1986.

5. D. G. Cain. "SMART: An Expert System Development Toolkit for the IBM-PC." EPRI NP-5645-CCM, August 1988.

6. R. Touchton, N. Subramanyan, A. Lane, P. Hariharasubramanian. "Nuclear Power Application of NASA Control and Diagnostics Technology." EPRI Report to be published.

7. J. R. Jamieson. "Knowledge-Based Automatic Test Equipment." Proc. Southcon/86, March 1986.

8. B. Hayes-Roth. "A Blackboard Architecture for Control." Artificial Intelligence, 26:251-321, 1985.

9. C. Horne. "A Brief Description of TRESCL-86." EPRI Internal Draft Report, November 1986.

10. W. Petrick and K. B. Ng. "Emergency Operating Procedures Tracking System." EPRI NP-5250M, June 1987.

11. R. A. Touchton et al. "Reactor Emergency Action Level Monitor Expert System Prototype." EPRI NP-5719, September 1988.

12. "Keystone Expert System Development Delivery Environment User's Manual." Technology Applications Incorporated, 1987.

13. A. Deardorff. "LIFEX: Expert System for Providing Guidance on Material Degradation Mechanisms." EPRI NP-6058, March 1989.

14. J. A. Naser, R. W. Colley, J. Gaiser, T. Brookmire and S. Engle. "A Fuel Insert Shuffle Planner Expert System." Seminar on Expert Systems Applications in Power Plants, May 1987.

15. B. M. Rothleder, W. S. Faught, G. R. Poetschat, and W. J. Eich. "The Potential for Expert System Support in Solving the PWR Fuel Shuffling Problem." International Topical Meeting on Advances in Reactor Physics, Paris, April 1987.

16. D. J. Finnicum and R. C. Webber "Thermal Performance Diagnostic Manual for Nuclear Power Plants." EPRI NP-4990, April 1987.

17. K. Yoshida and J. A. Naser. "A Proof-of-Concept Transient Diagnostic Expert System for BWRs." EPRI NP-5827-SR, May 1988.

18. D. G. Cain. "BWR Shutdown Analyzer Using Artificial Intelligence Techniques." EPRI NP-4139-SR, July 1985.

19. E. H. Groundwater, M. L. Donnell, and M. A. Archer. "Approaches to the Verification and Validation of Expert Systems for Nuclear Power Plants." EPRI NP-5236, July 1987.

20. D. B. Kirk and A. E. Murray. "Verification and Validation of Expert Systems for Nuclear Power Plant Applications." EPRI NP-5978, August 1988.

21. E. H. Groundwater, D. B. Kirk, A. E. Murray, E. H. Stottler and D. D. Dodd. "The EPRI Knowledge Acquisition Workshop Handbook." EPRI NP-6240, February 1989.

22. R. Pack, R. Lazar and R. Schmidt. "Expert Training Systems to Enhance Plant Operations Decision Making." EPRI NP-6379, April 1989.

Fossil Power Plant Applications of Expert Systems: An EPRI Perspective*

L. JAMES VALVERDE A., JR., STEPHEN M. GEHL, ANTHONY F. ARMOR, JOHN R. SCHEIBEL, and S. MURTHY DIVAKARUNI

Abstract

During the past decade, the field of artificial intelligence (AI) has witnessed tremendous growth. In particular, knowledge-based expert systems have quickly come to fore as one of the fastest growing subfields of AI. In this paper we discuss the role of expert systems in the electric power industry, with particular emphasis on six fossil power plant applications currently under development by the Electric Power Research Institute.

1. Introduction

Confronted with issues such as rising fuel costs, aging power plants, and a fluctuating economy, the electric power industry faces many challenges in the coming decades. Faced with these uncertainties, electric utilities are finding it increasingly difficult to balance economic and environmental goals, while concomitantly planning for anticipated demand growth. Because of the large financial risks associated with the construction of new power plants, many utilities have decided to postpone adding new generating capacity. This strategy places the burden of providing needed generation upon existing power plants and, perhaps, independent power producers. A major challenge, then, to American utilities lies in producing sufficient amounts of low-cost electricity with the currently installed capacity [1].

In order to meet this challenge, electric utilities are seeking ways to improve overall plant performance. The Electric Power Research Institute (EPRI) has, in recent years, actively pursued research and development in areas specifically aimed at improving net output, plant availability, plant efficiency, and operating flexibility. The phenomenological complexities inherent to these parameters are such that a great deal of domain-specific knowledge and information is needed in order to effectively enhance overall system performance. Because of their limited ability to incorporate both symbolic and numerical information, traditional computational approaches to these problems have met with marginal success. As an alternative to these approaches, Artificial Intelligence (AI) methods · which are better able to process symbolic (i.e., nonnumeric) information than traditional computing methods -- have begun to gain increased use and acceptance within the electric power industry.

---

*Electric Power Research Institute, 3412 Hillview Avenue, Palo Alto, California 94303.

During the past decade, the field of AI has witnessed tremendous growth. In particular, knowledge-based expert systems (ES) -- systems that are able to process the knowledge and information of human experts in a given domain -- have come to fore as one of the fastest growing subfields of AI. On a fundamental level, ES can, to varying degrees, embody certain aspects that are intrinsic to human expertise. For example, human experts are able to apply various types of knowledge and information over a broad range of applications; consequently, they are able to make *effective* and *efficient* use of their knowledge. In a similar fashion, ES are able to incorporate knowledge and information from multiple sources. By combining this attribute with the high speed of modern computing equipment, ES can quickly process knowledge and information that is particular to a specific task or problem. Human experts are also characterized by their ability to explain, in most cases, the specific lines of reasoning used to solve a particular problem. Using what are called *backward chaining* techniques -- techniques that begin with the solution to a problem and work backwards through the lines of reasoning used to arrive at that solution -- ES are able to provide the logic or reasoning behind a given solution. To varying degrees, then, ES are capable of embodying those traits that we normally associate with human expertise.

Recognizing the potential for ES, EPRI has, in recent years, taken measures to advance the implementation of ES technology throughout the electric utility industry. In this paper we discuss the role of ES in the electric power industry, with particular emphasis on fossil power plant applications. In Section 2, we begin our discussion by identifying two fossil power plant application areas that stand to benefit most from ES and AI-based approaches to problem solving. Next, in Section 3, we review current EPRI research and development in six fossil power plant applications of ES, covering such areas as heat rate degradation analysis, feedwater heater and condenser problem detection, boiler tube failure analysis, and plant modifications. In Section 4, we conclude our discussion with an assessment of the role of expert systems and artificial intelligence in the electric power industry, as well as speculate on the potential impact that ES technology can have in meeting the nation's present and future energy needs.

2. Fossil Power Plant Applications of Expert Systems

In recent years, electric utilities have begun to place considerable emphasis on enhancing certain aspects of plant performance, particularly heat rate improvement and unit availability. In application areas such as mechanical diagnostics, plant monitoring and control, maintenance, failure analysis, construction, coal quality impacts, and environmental controls operations, ES are meeting with acceptance and success [3,9,10,11].

A number of factors must be taken into consideration when identifying potential fossil power plant applications of ES. The first consideration is fundamental to the design of any ES, namely, applications should be sought in areas where there exists sufficient expert knowledge. Perhaps equally important, the application should have the potential for significantly enhancing the operation of fossil power plants. Moreover, given that human expertise is, in many respects, a valuable commodity, it is desirable to seek applications where human expertise is expensive or scarce. In this light, prospective fossil power plant applications of ES applications should, in so far as possible, possess the following general attributes:

- The candidate application addresses a genuine power plant problem;
- The candidate application requires expertise that may be expensive or in short supply;
- The common forms and recurring structures in the problem domain of interest are best approached from a *heuristic* vantage point, rather than a numerically oriented one;
- Sufficient knowledge exists and is readily available to solve the problems that are particular to the domain of interest;
- The use of ES technology is expected to result in improvements in performance parameters that would not otherwise be attainable by traditional computational approaches;
- The required level of expertise and modeling for the system is nominally within the existing state-of-the-art for ES.

In addition to the above desiderata, it is important to give thorough consideration to how electric utilities will initially perceive ES technology; early failures can cast doubt, while dramatization of successes can overstate the true capabilities of the technology. Given that AI and ES are relatively new technologies to the utility industry, it is important to minimize any possible misrepresentations of the technology and its potential applicability. With this understanding, the initial applications of ES within a utility setting should have a measurable impact upon their intended applications; ideally, it is also desirable that these benefits be realizable within a relatively short period of time.

Working with utility representatives, vendors, and consultants, EPRI recently published an R&D plan [4] for fossil power plant applications of ES. In this report, two application areas are identified as having a high degree of user interest, as well as having the potential for expedient adoption and use within the industry: 1) plant operations; and 2) equipment diagnostics. In both of these application areas, domain-specific and plant-specific knowledge and information can be used to enhance unit performance and availability, and to identify developing mechanical problems.

3. EPRI Fossil Power Plant Expert Systems

The Fossil Power Plants Department at EPRI is currently developing six fossil power plant expert systems. Working with technical experts in the utility industry, these systems are being developed and tested in an off-line mode; after this first phase of development, several of these systems will be installed on-line in power plant control rooms, where they will undergo further validation and verification. The six projects are as follows:

- Boiler Tube Failure Diagnosis System;
- Electrical Generator Monitoring System;
- Turbine Condition Monitoring System;
- Heat Rate Degradation Advisor;
- Condenser and Feedwater Heater Advisors;
- Plant Modification Advisor.

3.1 BOILER TUBE FAILURE DIAGNOSIS SYSTEM

Boiler tube failures are the leading cause of availability losses in U. S. fossil power plants. Each year, the industry averages nearly 4% lost availability in large fossil plants due to boiler tube failures. The causes of most of these failures are understood in sufficient detail to allow the specification of operating practices and plant modifications to minimize the occurrence of future failures. In this regard, EPRI has developed a comprehensive program for reducing boiler tube failures, which is currently being demonstrated at a group of 16 utilities; by implementing this program, these utilities have achieved substantial reductions in availability losses due to boiler tube failures.

3.1.1 Use of Expert Systems in Reducing Boiler Tube Failures

A key aspect of boiler tube failure reduction is the need for determining the cause of each failure, so that effective corrective and preventive measures can be taken. Several utilities in the EPRI demonstration project have used an ES, based on the EPRI Manual for Investigation and Correction of Boiler Tube Failures [7], to help diagnose failure causes [8]. The ES, called ESCARTA, asks the user a series of questions about the location and appearance of the failed tube and any potential initiating events. The responses to these questions are used in a backward chaining procedure to determine the likely cause of failure. After identifying the likely failure mechanism, the ES then recommends corrective actions to prevent future failures.

The overall structure and functions of ESCARTA are shown in Figure 1. The main menu of the program provides access to a failure diagnosis module, a data base on tube failures, a module containing extensive information on the 22 possible failure mechanisms, and a data base on tube dimensions and specifications. Since the failure mechanism information module is keyed to the results of a failure diagnosis, at the conclusion of a session with this ES, the user can access information on repair and inspection procedures, root cause analysis, and corrective action that is specific to the specific failure mechanism. The mechanism-specific data base supplements the information contained in [7] with information drawn from the EPRI Fossil-Fired Boiler Tube Inspection Guidelines [5], as well as results from ongoing EPRI projects in the boiler inspection and maintenance area. All of the data base modules can be easily modified by the user, for example, to add information on the particular repair procedures used by the individual utility, or to reference reports describing similar failures previously experienced at the plant. The ability to integrate data from several sources and provide the user with a concise summary of relevant facts and recommenda-

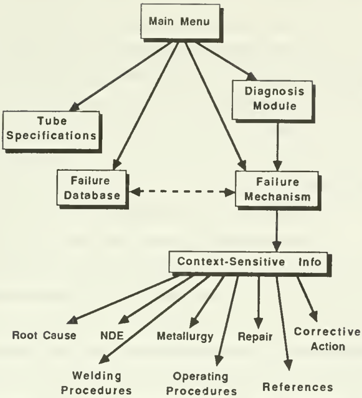


Figure 1. ESCARTA Structure and Functions

tions in the form of context-sensitive information is one of the advantages most often cited by users of this ES.

This ES has three broad application areas: (1) preliminary diagnosis of failure mechanism and probable root causes at the time of a failure; (2) quality control of the diagnosis process; and (3) training of plant personnel. When used for preliminary diagnoses, plant maintenance personnel can obtain rapid feedback on the mechanism and probable root cause of a failure. In practice, the results of the preliminary diagnosis are then conveyed to the central engineering staff and metallurgical experts for confirmation and to guide the planning of a detailed post mortem examination of the failed tube. By having access to a preliminary failure diagnosis at the time a failure occurs, the plant staff will frequently be able to select the proper repair procedure, return the plant to service with minimum delay, and in some cases, take immediate corrective action to prevent recurrence. Because it fosters the adoption of a precise vocabulary for describing failures and their effects, ESCARTA can also improve communications between plant personnel and general office staff.

The quality control function of ESCARTA is derived from its consistent automation of the diagnosis process. Questions are always asked in the same order (given the same responses), and relevant questions are never omitted. Consequently, utilities can use the diagnosis module to assure that all promising lines of reasoning are explored, thus minimizing possible misinterpretations of key symptoms.

In a training environment, this ES allows maintenance personnel to participate directly in root-cause analysis procedures, thus familiarizing them with the methods by which events, locations, and failure appearances are used in root cause analysis. Frequent references to [5, 7] and other references sources direct users to relevant information and, in the process, teach them to look for significant indicators in similar future situations. Experience with utility users of ESCARTA indicates that it teaches them to ask the key questions that are needed to identify root causes and distinguish superficially similar failure modes.

3.1.2 Boiler Maintenance Workstation

EPRI is expanding the applications of ES in the boiler availability area by developing a Boiler Maintenance Workstation (BMW). The objective of this project is to improve the accessibility and increase utility usage of EPRI products in the areas of boiler maintenance and availability. In its initial form, the workstation will include a version of ESCARTA for failure diagnosis and other EPRI software products in the areas of boiler inspection, maintenance, and life assessment. Workstation modules will analyze and display wall thickness data for water-wall tubes, predict the optimum time for inspections and tube replacement, perform creep life calculations for superheater and reheater tubes, and evaluate the remaining life of dissimilar metal welds in boiler tubes. As an aid in the failure diagnosis process, the workstation can be coupled to an optional 35mm slide projection or video disk system for displaying images of failed tubes. This will allow utilities to add photos of their own failures, which may differ from the textbook examples contained in [7].

The workstation is designed to run on Intel 80286- and 80386-based microcomputers. A typical utility implementation will have workstations at the general engineering offices and at every fossil steam plant on the system. Ideally, the workstations at the power plants will be electronically connected with the engineering office system so that the "master" version of the data base modules will be updated as soon as new information becomes available. EPRI plans to sponsor a demonstration of the BMW at a group of host utilities. The utilities participating in the demonstration will evaluate the workstation over a six-month period, report on their experiences, make recommendations for modifications and additions to the workstation, and document the benefits of using the BMW in their boiler maintenance programs. The results of these utility demonstrations will be available in late 1990.

The BMW is one of the applications currently under development as part of the EPRIGEMS program, a new program at EPRI that endeavors to use ES as a means of effecting technology transfer of EPRI R&D results [2]. The EPRIGEMS user interface will make the BMW and its components accessible to a wider utility audience. In addition, the modular grouping of the component programs in the workstation will facilitate information transfer among the programs. The boiler tube failure diagnosis module is the only ES incorporated into the first version of the BMW, which is scheduled for release in the fourth quarter of 1989. Subsequent versions of the program will make expanded use of AI techniques to guide the user through the applications of the various component programs.

3.2 ELECTRICAL GENERATOR MONITORING SYSTEM

The reliability of turbine generators is critical to fossil power plant reliability and operation. In order to minimize prolonged generator outages, it is important to receive early warning of machine problems before failure. Recognizing the growing need for such capabilities, work is currently under way at EPRI to develop an on-line generator monitoring system. This system will correlate available generator diagnostic information obtained from sensors to advise operations personnel of developing generator problems. Having identified a potential generator problem, the monitoring system then makes relevant recommendations for corrective action.

At the core of this ES is the knowledge base and the inference engine. The knowledge base consists of an extensive set of rules, elicited from experts in the field, that identify the likely sources of trouble in the generator. The inference engine then uses this stored knowledge and information to analyze sensor input and offer solutions and recommendations relevant to the problem at hand.

The required flow of information in the Electrical Generator Monitoring System presents many technical challenges. First, data from machine sensors enters a data collection subsystem, and then enters a status evaluation module, which examines the data for trends that may be indicative of problematic phenomena. When such phenomena is detected, the flow of control is then passed to the inference engine, which draws upon the knowledge base to prescribe a relevant course of action for the observed phenomena. The monitoring system will also qualify its recommendation by providing a confidence level, a level of urgency, and a measure of severity. This type of information will be extremely helpful to the operator in judging the scope and immediacy of the current problem.

An important feature of this system is the installation advisor, which allows for the customization of the system to the particular generating unit that it will be used with. This customization allows plant engineers to incorporate important plant-specific details of the generator and its sensors, as well as the operating policies of the utility.

The first Electrical Generator Monitoring System will be installed on-line at the Nanticoke Station of Ontario Hydro, the prime contractor, by the end of 1989. The second system will be installed in 1990 at the Oswego Station of the Niagara Mohawk Power Corporation.

3.3 TURBINE CONDITION MONITORING SYSTEM

Because of their ability to integrate both numeric and symbolic information, ES are well suited to the task of complex diagnostic process monitoring, where many fault types and multiple symptoms must be considered. In diagnostic monitoring of steam turbines, vibration signatures can be ambiguous and equipment dependent. This, of course, makes specific fault definition a complex and inherently uncertain task. For example, a vibration with a periodicity equal to the running speed may be caused by a change in unbalance force, system stiffness, or system damping. On the other hand, a vibration at twice the running speed may be caused by a change in rotor or bearing stiffness, or perhaps by misalignment of the rotor at the bearings. To mistake high vibration caused by a rotor crack for unbalance or misalignment of the turbine rotor can be a costly error.

Vibration and acoustic signature data from operating turbines are analyzed using various signal processing techniques that help discriminate between different fault types. In addition to signature data, other types of data may be required. For example, rotor position, bearing temperature, or performance data may reveal problematic phenomena that requires attention. An ES provides an ideal framework from which to perform diagnostic evaluations, for it can draw upon a range of sensor data, calculated values obtained from physical models, and information contained in data bases.

At the Florida Power & Light Port Everglades Station, EPRI and General Electric are currently demonstrating a Turbine Condition Monitoring System [9]. This ES acquires on-line turbine generator condition data directly from a microprocessor-based vibration signature analysis monitor. Vibration, temperature, shaft position, and phase angle are all monitored during steady-state and coast-down operation. A minicomputer then performs the data collection, processing, and numerical analyses, while a PC performs the symbolic ES diagnosis.

The knowledge base of the Turbine Condition Monitoring System contains about 150 rules and diagnostic strategies directed towards seven major fault types. Table 1 lists the major fault types that can then be attributed to twenty-six specific mechanical failure causes. For example, the system can determine if misalignment can be attributed to, among other things, the bearing or the coupling. A typical diagnostic rule checks whether a particular condition is true or false. If the condition is true, then a weighting factor -- a measure of the condition's significance as a fault symptom -- is applied.

MAJOR FAULTS	SPECIFIC FAULTS					
UNBALANCE	LOSS OF MASS	EROSION	1ST STAGE EROSION	STOP VALVE BYPASS FAILURE	BEARING WEAR	
RUB	RADIAL	REGULAR RADIAL	CARBONIZATION RADIAL	PACKING RUB	AXIAL RUB	
BOW	WATER INDUCTION	THERMAL SENSITIVITY	RESIDUAL BOW			
MISALIGNMENT	BEARING	BEARING VERTICAL	BEARING ANGULAR	COUPLING	PARALLEL COUPLING	ANGULAR COUPLING
WHIRL	OIL	STEAM	RESONANCE			
MOUNTING	LOOSE BOLTS	EXCESSIVE CLEARANCE	BORE PLUG			
ROTOR CRACK	TRANSVERSE					

Table 1. Major Fault Types

3.3.1 Misalignment Diagnostics

To illustrate the logic used in the Turbine Condition Monitoring System, consider, for example, the shaft-bearing misalignment fault diagnosis process. This process follows four steps:

1. Sensor data is collected once per hour and entered into a data base. Bearing, coupling, axial positions, bearing metal temperature, and displacement data are stored by time, load, and steam temperature.

2. The numeric sensor data is then used to respond to system queries in the form of true or false statements. For example, a bearing metal thermocouple reading greater than 15° F is defined as a 'true' state for the condition 'abnormal metal temperature'. In a similar fashion, sensor data relating to vibration, shaft position, and bearing temperature is used to describe the various physical states of the system.

3. The symbolic facts are used to respond to rule base questions shown in Table 2. Screening rules determine the most probable major faults, followed by a general and then specific fault analysis. If, for example, the axial position or the bearing metal temperature is abnormal, then the general and the specific case for misalignment is investigated. Each rule found to be true is assigned a weighting factor proportional to its importance. A total weight for each investigated major fault is then determined.

4. Major faults are ordered from highest to lowest nonzero total weight. The major fault is then listed with the specific fault determination. For example, referring back to Table 1, a major fault could be 'whirl', and the specific fault determination could be either 'oil', 'steam', or 'resonance'.

	MISALIGNMENT RULES	TRUE	FALSE	FIRE	VALUE
MAJOR FAULT SCREENING (PARTIAL LISTING)					
1	IF abnormal D C position THEN investigate MISALIGN	T		•	
2	IF abnormal bearing metal temperature THEN investigate MISALIGN		F		
GENERAL MISALIGNMENT					
3	IF 1/rev phase is steady and 2/rev phase changes THEN add W3	T		•	W 3
4	IF bearing metal temperature is abnormal and D C position is		F		
	abnormal THEN add W4				
5	IF there is a significant difference between adjacent bearings'	T		•	W 5
	metal temperatures or orbits or D C position THEN add W5				
6	IF any coupling D C positions are abnormal THEN add W6	T		•	W 6
7	IF axial metal temperature is abnormal THEN add W7		F		
8	IF axial D C position is abnormal THEN add W8		F		
ANGULAR COUPLING MISALIGNMENT					
9	IF axial metal temperature is abnormal THEN add W9		F		
PARALLEL COUPLING MISALIGNMENT					
1 0	IF relative changes in D C position and phase occur between	T		•	W 1 0
	adjacent coupling probes occur THEN add W10				
VERTICAL BEARING MISALIGNMENT					
1 1	IF 1/rev and sub-synchronous is abnormal THEN add W11		F		

Table 2: Misalignment Rules. Sensor data for vibration, rotor position, and bearing metal temperature is used to assign truth values to each possible system state. The rules in this table are arranged so as to first determine the most likely major faults, and then proceed with a more detailed analysis to confirm the fault type and its mechanical cause.

Work in this area is continuing to expand the rule base to include additional faults and fault symptoms.

The automated analysis and interpretation of sensor data that the Turbine Condition Monitoring System provides holds promise to improve the effectiveness of both periodic and continuous condition monitoring programs. By approaching this problem from an ES vantage point, large amounts of data collected from periodic machinery surveillance programs using portable vibration spectral collectors, as well as from continuous monitoring turbine supervisory instrumentation, can be more efficiently screened and related to performance and maintenance data. Since an ES can readily supply routine fault analysis, vibration and equipment specialists will be better able to focus on events that are likely to warrant attention by plant engineers.

3.4 HEAT RATE DEGRADATION ADVISOR

EPRI is developing and demonstrating an ES to help utility operators and engineers diagnose and correct the conditions that lead to heat rate losses in fossil power plants. The objectives of this project are to enable utilities to achieve a measurable improvement in heat rate through improved response to both major and minor changes in plant operating conditions, while providing sufficient flexibility of design to facilitate widespread implementation throughout the industry.

Historically, many utilities have monitored heat rate on a monthly basis by the ratio of total fuel consumption to total gross generation. This measure of heat rate is most useful as a rough estimate of operating costs, but is not suitable for diagnosing problems or trending plant perfor- mance. Another common practice is periodic performance testing using on-line measurements of temperatures, flows, and pressures to determine the efficiency of key plant components. Periodic performance testing effectively indicates heat rate problems that require corrective actions, but, because of the extended intervals between such tests, heat rate degradation frequently goes undetected for long periods of time. Periodic performance testing does not provide either plant operators or performance engineers with the information that is needed to improve or maintain heat rate as operating conditions change.

An ES capable of accurately diagnosing heat rate losses in a time frame that allows rapid identification and correction of the underlying problem must be based on a thorough understanding of the factors that affect plant performance. Such a system must also have access to on-line performance information. Previous attempts to develop heat rate expert systems have been specific to a particular power plant, and have not been generally applicable across the industry. EPRI has adopted the approach of designing a heat rate ES for maximum flexibility, so that it will be applicable to plants of differing design with different levels of performance monitoring instrumentation. The information on plant performance issues in the Heat Rate Degradation Advisor will come, in part, from the Heat Rate Improvement Guidelines for Existing Fossil Plants [6], which outlines an approach for identifying the root causes of heat rate degradation and implementing corrective actions. These guidelines include a set of heat rate logic trees that are used to help diagnose the likely source of heat rate losses. As exemplified in Figure 2, a logic tree begins with a statement of the problem being addressed, identifies all the failure modes associated with that problem, reduces the failure modes to the underlying root causes, and identifies the information needed to verify the root causes. The logic trees are designed to be applicable to a wide variety of plant designs, and the information in [6] will be supplemented with analytical relationships and heuristic knowledge to enable the interpretation of on-line data. The result will be a set of diagnostic rules that will cover nearly all plant designs and modes of operation.

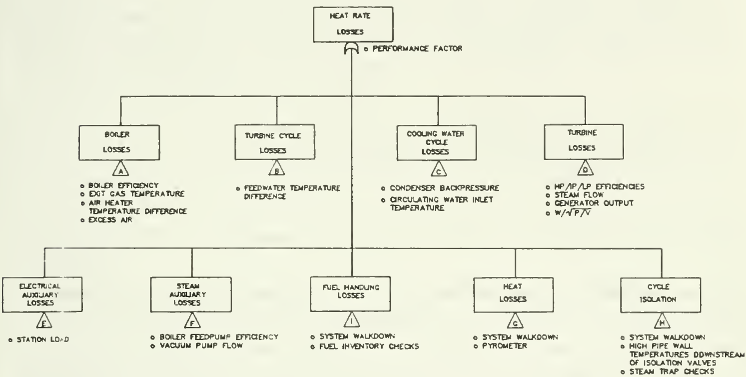


Figure 2: Top-Level Heat Rate Logic Tree. This logic tree shows broad categories of heat rate losses. Subsequent logic trees in this series give progressively more detail on the causes of plant performance problems.

Figure 3 provides a block diagram of the Heat Rate Degradation Advisor. The ES will be designed to accept input from three major sources: (1) sensor data currently logged by the plant computer; (2) data from sensors not coupled to the plant computer; and (3) manual input of off-line measurements and qualitative observations. Furthermore, the ES will be designed to accommodate differences in the numbers and types of sensors in each individual implementation. An important part of the system development will lie in determining the minimum set of sensors needed to get acceptably accurate diagnoses and recommendations, and the level of accuracy achievable with different levels of plant instrumentation. Figure 3 also shows that the Heat Rate Degradation Advisor will be designed to operate in conjunction with an existing on-line performance monitor. The system will also have internal performance calculation models for use in applications without a separate performance monitor.

BASIC INFORMATION FLOW DIAGRAM

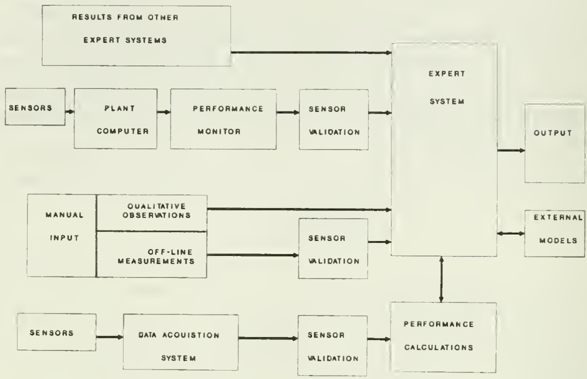


Figure 3. Heat Rate Degradation Expert System Advisor

The user interface of the Heat Rate Degradation Advisor will emphasize the needs of the plant operator. For example, extensive use will be made of graphic presentations of plant conditions, including significant deviations from optimal values. Presentation screens will include menus, graphics of individual components and systems, graphic illustrations of identified trends, text windows, and data tables. The user will also be able to access additional screens that contain the input data and logic used by the expert system to diagnose a particular condition. Recommendations of the system will be keyed to an extensive on-line data base of information on the correction and prevention of heat rate degradation. The data base will also contain citations to outside sources of information. In addition, the data base will be customizeable by the user to add plant- or utility-specific information.

The expert system development project is planned in two phases over a four-year period. Phase I (1989-1992) will consist of development and industry demonstration, and phase II (1992-1993) will consist of the commercialization activities.

3.5 CONDENSER AND FEEDWATER HEATER ADVISOR

Condensers and feedwater heaters (FWHs) are frequent sources of unit unavailability and heat rate losses. In the course of normal operation, FWHs are susceptible to a number of possible failure modes and performance problems, the most likely of which are tube failures. Other failure modes include adverse water chemistry conditions, plugged vents, and valve/controls failures. For condensers, tube bundle design problems, excessive air in-leakage, air removal equipment malfunction, circulating water system, and macro/micro fouling all contribute to condenser performance problems.

EPRI is developing expert systems to aid in diagnosing performance degradation and failures or malfunctions of condenser and feedwater heater systems. The overall structure of these expert systems will be similar to that of the heat rate degradation expert system described above. In particular, these systems will be able to accept manual input and data from the plant computer, as well as data from sensors that are not connected to the plant computer.

The initial focus of the FWH Advisor will be off-line fault diagnosis. Since most feedwater heater problems develop slowly, there is little benefit in having real-time data analysis capability for real time data analysis. This situation may change, however, particularly for plants that have installed on-line leak detection systems. For this reason, the feedwater heater expert system is being designed for easy modification to on-line data analysis.

In contrast to the FWH Advisor, the Condenser Advisor is being designed as an on-line system. By continuously monitoring plant performance parameters, the Condenser Advisor will, in many cases, be able to diagnose faults and prescribe corrective action before severe damage occurs to the unit. In addition, the on-line monitoring of performance degradation will allow for scheduling of maintenance activities. The Condenser Advisor will also work well in conjunction with planned on-line condenser maintenance activities, such as tube cleaning, targeted chlorination, and on-line tube leak plugging.

The development and demonstration of the condenser and feedwater expert systems will closely follow that of the Heat Rate Degradation Advisor development in 1989-1992.

3.6 PLANT MODIFICATION OPERATING SAVINGS

Changing industry and economic conditions are forcing utilities to reevaluate cost-minimizing operating practices of fossil power plants. Older plants were designed principally for single-shift, non-cycling operation, restricting the ability to economically dispatch these plants to meet fluctuating load conditions. Any modifications made to these plants to enhance low-load operating efficiency and/or cycling capability must be made on a cost-effective basis. In this regard, it is necessary to employ analytical models that can consistently and accurately estimate highly uncertain future benefits. Historically, stand-alone financial models have been unable to capture sufficient technical detail, while highly detailed engineering models have been unsuccessful in translating changes in technical specifications into financial impacts. Ideally, a robust evaluation methodology should combine the underlying technical knowledge of plant modifications with appropriate valuation models. EPRI is currently developing a system, the Plant Modification Operating Savings (PMOS) system, that seeks to combine these two approaches. PMOS differs slightly from the five ES described above in two ways:

- While most ES applications are designed to provide either *ad hoc* diagnosis or consultation of fossil power plant subsystems, PMOS was designed to provide insights into the *future* impacts of modifications on plant performance;

- The principal structure of PMOS is numeric rather than symbolic.

Although the ES paradigm is based, primarily, on heuristic approaches, some problems require additional analytic capability. Accurate estimates of plant modification benefits require an assessment of optimal plant operation on a before/after basis over a complete time horizon. The preferred method for this type of assessment is based on dynamic programming (DP), a mathematical technique for making a sequence of interrelated decisions. Without adequate formulation and bounding of the problem, however, the run-time of a standard DP algorithm can rise exponentially. PMOS uses a set of heuristics that combine knowledge of plant modification impacts and dynamic programming techniques that bound the estimation problem based on individual power plant characteristics.

As illustrated in Figure 4, PMOS consists of two related systems sharing central data storage and viewed by the user as a single, integrated system. The evaluation controller contains heuristics that bound the problem by determining appropriate procedures and parameters that are unique to each modification. Given this formulation, the evaluation engine uses DP to perform an estimation of modification benefits for a given time period. The controller uses the engine iteratively to estimate the benefits for an entire time horizon as specified by the user. Operating and performance results (including estimated benefit/cost ratios) are ultimately delivered via reports and graphs.

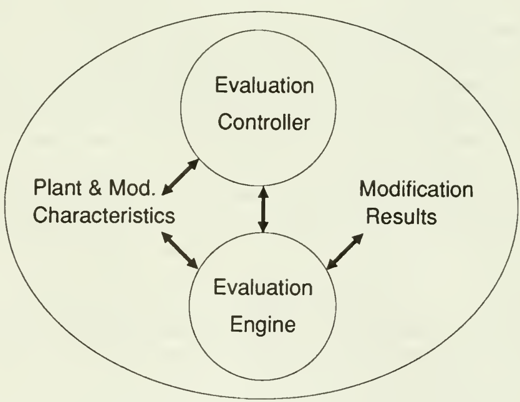


Figure 4. PMOS Structure

A prototype version of PMOS has been used to evaluate ten major fossil power plant modifications for the Duke Power Company. These modifications included:

- Heat rate improvements;
- Low load modifications;
- Variable pressure operation;
- Control system upgrade.

The formulation of PMOS provides the capability to evaluate any modification that can be characterized by an impact on any of the following plant cost and performance characteristics:

- Fuel costs and variable O&M costs;
- Loadings and heat rates;
- Ramping ability and associated fuel and stress costs;
- Start-up fuel and stress costs;
- Hot standby feasibility.

Enhancing ES technology and delivery systems with existing quantitative methods is a valuable combination. Advanced mathematical models require the type of control available under heuristic systems, while many quantitative tools require analytic models and technical knowledge bases a their core. PMOS demonstrates how these varied paradigms can be unified within a shell whose goal is financial valuation. Figure 5 illustrates the relationship between lower-level technical ES and analytic models with higher-level financial valuation systems. Integrating value models for all the principal components of a fossil power plant results in an integrated decision system whose use is more closely related to a utility's corporate goals and objectives.

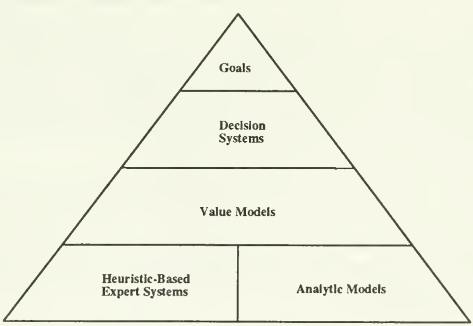


Figure 5. Intelligent Decision Systems

In light of the above discussion, two observations regarding the use of ES in the electric utility industry arise from the work performed thus far:

- Heuristic-based technical ES and quantitative or analytic models are not mutually exclusive;
- Some utility problems (e.g., plant modification) must contain both sets of tools, integrated within a financial valuation framework.

A production version of PMOS is currently under development and is scheduled for several utility applications during the summer and fall of 1989.

4. Conclusion

Electric utilities currently find themselves in an increasingly competitive and uncertain environment. Consequently, they must seek technological advances in areas that can minimize the costs of producing electricity. This objective can be realized in a number of ways, the most obvious of which is to improve the efficiency and reliability of the existing generating capacity. In this paper we have discussed how AI and ES technology is being used to help utilities achieve this goal.

The extent to which ES technology will impact the electric power industry is not yet known. Nevertheless, it is clear that there exist a number of application areas that can benefit from the unique capabilities that this technology provides. However, in spite of the initial successes that the utility industry has had in applying ES technology, it is important to understand the current limits of the technology. In recent years, AI researchers interested in developing a general, unified approach to ES design have begun to examine formal models of knowledge and reasoning in order to better understand how to acquire and represent the deep knowledge that characterizes much of human expertise. A major problem in transferring knowledge from human to machine stems from the need to translate human knowledge into computable formalisms. Of course, this problem is further-complicated by the fact that much of the knowledge that a human expert uses is characterized by uncertainty. Consequently,the value of ES to practicing engineers will increase as improved mathematical methods for handling uncertainty are developed. In addition, continued developments in theoretical structures for knowledge acquisition and knowledge representation are anticipated, thus facilitating the implementation of complex engineering applications.

EPRI's initial focus on ES development has been in technical domains where extensive research and development has been conducted; consequently, knowledge representation and uncertainty management has been relatively straightforward. The experience gained through utility implementations of these ES will provide the basis for the development of systems capable of addressing a broad class of engineering applications.

Acknowledgement

The authors wish to thank D. Cain, Peter C. Borocz, Salim J. Jabbour, and J White for their insightful comments and suggestions on earlier drafts of this paper.

References

1. Armor, A. F., et al., "Expert Systems for Power Plants: The Opportunities," American Power Conference, Chicago, IL, April, 1988.

2. Cain, D., et al., "EPRIGEMS: Expert Systems for Technology Transfer," EPRI Conference on Expert Systems Applications for the Electric Power Industry, Orlando, FL, June 5-8, 1989.

3. Dohner, C. V. and S. J. Acierno, "Gas Turbine Expert System," American Power Conference, Chicago, IL, April 1988.

4. *Expert Systems Technology Assessment and R&D Plan for Fossil Plant Applications*, Electric Power Research Institute, Palo Alto, CA, January, 1988.

5. *Fossil-Fired Boiler Tube Inspection, Volume 1: Nondestructive Testing Guidelines*, Electric Power Research Institute Report CS-4633, August, 1986.

6. *Heat-Rate Improvement Guidelines for Existing Fossil Plants*, Electric Power Research Institute Report CS-4554, May, 1986.

7. *Manual for Investigation and Correction of Boiler Tube Failures*, Electric Power Research Institute Report CS-3945, April, 1985.

8. Nichols, V., et al., "Utility Experience with ESCARTA: An Expert System and Data Base for Preventing Boiler Tube Failures," Second EPRI Fossil Plant Inspections Conference, San Antonio, Texas, November 29 - December 1, 1988.

9. Scheibel, J. R. and C. J. Carlson, "On Expert System Based On-Line Rotor Crack Monitor for Utility Steam Turbines," Second International Conference on Condition Monitoring, London, England, May, 1988.

10. Sopocy, D. M., et al., "Expert Systems to Improve Plant Operability," American Power Conference, Chicago, IL, April, 1988.

11. Thompson, T. and M. Coffman, "TU Electric Experience with On-Line Generator Monitoring and Diagnostics," American Power Conference, Chicago, IL, April, 1988.

Review of Expert Systems in Power System Operations Electrical Systems Division

DAVID CURTICE
Electric Power Research Institute
Electrical Systems Division
3412 Hillview Avenue
Palo Alto, California 94304, USA

ABSTRACT

This paper reviews some of the expert system research projects of the Electrical Systems Division of EPRI. It presents the results of expert systems developed for power system operations.

To date, two of the three expert systems developed for system operations are currently being evaluated by system dispatchers. Plans call for developing two more expert systems for alarm processing and scheduling for demand-side management programs.

INTRODUCTION

EPRI believes there is a significant potential for expert systems to aid power system dispatchers in a number of procedures that are frequently encountered in operating power systems. Although the performance and the speed with which expert systems will find their way into every day application are easily over-stated, research-to-date confirms that the basic premises of applying expert systems for power system operations tasks are, indeed, valid.

BACKGROUND

Power system dispatchers continuously monitor and supervise the power system.

They normally implement actions that are for the most part preplanned. These preplanned actions are based on operations studies of the system performed in planning and operations planning that consider (at least ideally) all the likely planned and forced outages.

Even when the power system is in a normal state, however, conditions are not predictable. System dispatchers must constantly deal with loads that depart from estimates, unavailability of planned for generating units and innumerable other contingencies.

With the increasing capability of energy management systems, system dispatchers are receiving a formidable volume of numerical data that must be routinely examined and interpreted to determine which actions should be taken.

System dispatchers are becoming overloaded with data. Interpretive programs are needed to evaluate data and tell the operator things that he/she needs to know.

The system dispatcher is inundated with alarms when a significant upset occurs. While progress has been made in giving priority to certain classes of alarms, what is needed is a system sufficiently "smart" to identify the initiating contingency and/or that part of the network which should receive the dispatchers first attention.

Expert systems should help the dispatcher to diagnosis system problems, point out the right direction and suggest alternative actions to deal with the problem. And provide the dispatcher with information that predicts the results of his actions before they are implemented in the real system.

System dispatchers are responsible for maintaining a match between generation and load, ensuring that equipment operates economically within allowable bounds. In managing a network emergency, dispatchers must restore normal operation while avoiding equipment damage and loss of service to customers. Expert systems incorporating the expertise of numerous personnel may help to control emergencies more effectively than a single dispatcher, thereby improving the utility's service to customers.

Dispatchers must convert great quantities of numerical data into information for assessing power system performance. With energy management systems now being equipped to handle 600 alarms per minute and up to 2000 in 15 seconds during emergency conditions - dispatchers experience data overload, which might lead to severe consequences in emergencies. Artificial intelligence (AI) technologies - expert systems in particular - have the potential for converting voluminous data into usable information. Ultimately, these technologies could diagnose power system problems, provide operators with analysis of system malfunctions, and suggest preventive or corrective actions.

OVERVIEW OF RESEARCH AND DEVELOPMENT PROJECTS

Research project RP1999-7 was developed to identify and evaluate uses for AI technologies in power system operations and to demonstrate the potential of two such technologies--expert systems and symbolic programming--for power system control.

Investigators collaborated with Allegheny Power System engineers to identify 16 potential applications of AI in power system operations. They collected data to determine whether using AI in those applications would be feasible and, if so, whether it would significantly improve existing problem-solving strategies. They also developed a system for integrating numerical and symbolic processing and two AI-based programs. To provide information for planning projects that would not duplicate work already under way, they identified utility-related AI research being conducted by other R&D groups (1).

A demonstration prototype, containing about 600 rules and written in OPS-5 running on a DEC VAX 11/780 computer, was developed for troubleshooting transmission relays and breakers.

Results of the study provided a foundation for future work. Of the 16 AI applications reviewed, only one - contingency selection-security assessment-met all of the researcher's feasibility criteria. This application was recommended for further study. The other applications - alarm processing, economic control and preventive control - met most of the criteria. The researchers suggested that these applications also be investigated (1).

The demonstration phase of the study produced two programs that illustrate the potential benefits and current limitations of AI for power system applications. One program uses a variety of relay models and coordination modes to simulate power system protection schemes. The other, a program for diagnosing faults, identifies disturbances or equipment malfunctions that initiate changes in network configurations. A system was also developed to link symbolic and numerical programming languages.

This study constituted our first comprehensive investigation of how expert systems might be applied in power system operations and showed that such systems do hold promise for solving long-standing power system analysis problems. The small number of value, large scale applications found to be feasible, however, suggests that utilities should use caution in estimating the potential of AI and that the use of expert system for solving such problems as unit commitment, maintenance scheduling and fuel scheduling should be examined more thoroughly. Moreover, the large number of rules (600) used to develop two very simple AI-based demonstration programs raised questions about the performance requirements of more complex programs and whether the logic segments of one program can be transferred to another.

Research project RP1999-9 was developed based on the results of RP1999-7. The objective was to build a prototype expert system for emergency control of power stations. Specifically, this project has developed a prototype expert system for Customer Restoration and Fault Testing (CRAFT) to assist system dispatchers perform on-line analysis to locate faults causing transmission line outages. The CRAFT system is the first step in a broader effort to build an experimental expert system for the emergency control of power systems (2).

The project team first interviewed Puget Power System dispatchers, who described the procedures and reasoning they use to solve problems manually. They used this expertise to develop approximately 300 rules for fault isolation and service restoration. They then incorporated these rules into the prototype CRAFT expert system to serve as a dispatcher's aid and demonstrate the proposed actions, they revised the rules to handle new situations and give more-accurate responses. Finally, the team developed a plan to implement such a system in an actual control center. They studied two feasible approaches. An appended approach would put the expert system on a separate computer, linked to the center computer with minimal disruption of its operation and displays. An embedded approach would integrate the expert system into the central computer, providing quicker responses than the appended approach (2).

One goal of EPRI's power system planning and operations research is to automate those tasks best handled by computers, thereby helping member utilities plan and operate their power systems more efficiently. The key to this goal is implementation of expert systems to aid and interact with dispatchers. A host of tools is currently available to help dispatchers with normal on-line network operation, and work continues to improve these tools. Once the power system transits to an emergency state, however, dispatchers and operators have far fewer tools to help steer the system out of trouble. In addition, utility experts are not always available for consultation. By providing efficient assessment of system conditions and suggested remedies based on utility philosophy and judgement, expert systems can quickly provide the operator with options.

EPRI, Puget Power and the National Science Foundation are cosponsoring continuing EPRI project RP1999-9 to implement CRAFT on-line at Puget Power. In addition to reporting the experience of Puget Power system dispatchers, this project will further study the embedded and appended implementation approaches and develop other areas in which expert systems can assist dispatchers, such as fuel allocation and use, voltage profile enhancement, and security analysis (4).

Research project RP2473-8 was developed to compare different languages used to implement expert systems. Two widely used computer languages, Program In Logic (PROLOG) and Official Production System (OPS), exist for developing expert systems. On a previous project, RP1999-7, a prototype expert system was developed for simulating the behavior of protection schemes in power systems. It was written in OPS-5 and performed adequately. This project undertook the task of translating from OPS-5 to PROLOG (3).

Subsequently, RP2473-8 developed a Volt/VAR dispatch system using PROLOG. It provided a simulation of the protection system and a realistic model of Union Electric Co.'s power system with a link to a FORTRAN power flow program to provide a simulation of the power system (5).

In applying expert systems to solve power system operation problems, PROLOG appeared to have an advantage over OPS, which starts with a set of known facts and searches for a conclusion based on these facts. PROLOG, on the other hand, begins with a goal and searches for facts to support that hypothesis. Because many power system algorithms employed by utilities are goal oriented, such as Volt/VAR dispatch, PROLOG might be a suitable choice for developing the expert system.

Recently, proposals were requested from selected bidders to develop, demonstrate and commercialize expert system for use in power system operations. Projects funded under this initiative consist of two phases. The first phase will develop several prototype expert systems for evaluation. The second phase will demonstrate and then commercialize the best prototypes from the first phase.

Several projects will be funded to develop a comprehensive package of expert systems for power system operations. To accomplish this goal, EPRI seeks to fund projects that will produce commercial expert systems. In general, these expert systems would have the following characteristics:

a) Relieve human expert of routine decision making.
b) Contain knowledge and data about the problem that is readily available.
c) Contain some information associated with the problem that is judgemental, i.e. based on experience gathered over the years by experts.
d) Based on problems that can be logically divided into stages.
e) Have outputs that can be evaluated.

At this stage, interest in expert systems focuses on those activities with the highest payback, such as:

a) Productivity improvements: human as well as machine productivity improvements.
b) Fuel expenditures.
c) Reliability: reliability and operating security.

Productivity and fuel expenditures currently dominate the industry's focus because utilities must remain the low-cost supplier of energy services. Reliability and power system security are very important but are more difficult to quantify in dollars.

ISSUES

The promise and potential contribution of expert systems could lead to prodigious achievements. Despite their limitations, expert systems do not tire, they don't forget, and they don't get emotional or frantic under stress. Their ability to recall vastly more encoded knowledge than any human can hold in memory is perhaps their strongest feature.

The challenge to EPRI's R&D projects is to integrate expert systems into an environment dominated by FORTRAN and the tightly coupled software and hardware used in energy management systems. And equally important is EPRI's goal of transferring expert system technology to its members.

Expert systems for power system operations must be developed with at least three (3) barriers recognized before the functional specifications are completed:

- Platform - integration with the energy management system (EMS) or linked to the EMS, e.g.; workstation,
- Uniqueness - are expert systems transferable from one utility's power system to the next,
- Maintenance - need for additional software and possible hardware expertise, and maintenance of rules or knowledge base.

While the problem of integration with the utility's EMS remains, there are new developments in workstations that maybe used as dispatcher consoles, providing that the workstation can emulate the EMS displays.

A major unresolved concern is the transferability of a developed expert system. Even if the software is not portable, we need to determine if the structure or the rules can be used by another utility.

Maintaining a new technology always increases the need for specialized expertise. Expert systems add another dimension to the problem of maintenance--knowledge base or rules maintenance. As new rules are developed, they must be entered, and checked to see if they are robust, or in conflict with existing rules, and if they are tautologies.

CONCLUSIONS

The Power System Planning and Operations program of the Electrical systems Division of EPRI has completed two (2) operating expert systems. Both are being evaluated by system dispatchers.

Several new projects have been started to develop prototypes for alarm processing, demand-side management, security enhancement, and optimization programs. These efforts are focused on high benefits to cost ratio applications.

REFERENCES

1. Artificial Intelligence Technologies for Power System Operations. Palo Alto, California. Electric Power Research Institute, EPRI EL-4323, January 1986.

2. Development of Expert Systems as On-Line Power System Operational Aids. Palo Alto, California. Electric Power Research Institute, EPRI EL-5635, February 1989.

3. Knowledge-Based System in Electric Utility Operation Using the PROLOG Language. Palo Alto, California. Electric Power Research Institute, EPRI EL-5666,

February 1988.

4. Experimental Implementation of an Expert System Dispatcher's Aid. RP2944-1 is an extension of (2) above. Final report due in October 1989.

5. Knowledge-Based System for Voltage and VAR Dispatch. RP2944-2 is an extension of (3) above. Final report due in June 1989.

The Need for Portable Expert Systems in the Workplace

CLARK DOHNER
Electric Power Research Institute
3412 Hillview Avenue
Palo Alto, California 94304, USA

Introduction

Trouble-shooting and diagnosing problems which arise in power plants can require expertise usually possessed by only a few experienced technicians. These experienced technicians could provide guidance to assist the less-experienced trouble-shooter, but they quite often are busy and not readily available. Expert knowledge can be extracted from these experienced trouble-shooters and implemented as rules in a computer-based system, called a knowledge-based or expert system. The expert system, then, can be used by the novice trouble-shooting technician - but only if he can access it in his workplace environment.

Background

In 1983, a project was initiated at EPRI to develop an expert system for trouble-shooting problems in gas turbine power plants. At that time, it was recognized that solution to the trouble-shooting problems contained two critical aspects:

1. The expert knowledge
2. User access to the expert knowledge (i.e., the man-machine interface)

Up to that time, most expert systems had been developed by knowledge-engineers who used higher level knowledge languages (such as LISP) for incorporating the rules they extracted from engineers, designers, and field personnel (i.e., the experts). These higher level knowledge-development tools usually resided on specialized computers or on main-frames. Thus, the ability to use this knowledge in the power plant workplace was severely limited and resulted in expert systems being used mostly in the fixed, office environment. Although the military, through DARPA (Defense Advanced Research Projects Agency), had funded some efforts in the direction of field-deployment of expert systems, there was no practical system available for taking a knowledge-base (including visual materials) to the power plant trouble-shooting workplace.

EPRI's project focused on these two crucial areas in an effort to:

1. Develop an expert system for performing a trouble-shooting task in a gas turbine power plant workplace by inexperienced technicians.

2. Develop a user interface which would:

    a. Allow the user to interrogate the expert system from the plant location where he necessarily must perform the trouble-shooting task

    b. Be easy-to-use

c. Provide the multimedia communication for assisting the user in performing this task, regardless of his preferences.

Solution

In developing the complete system, it was necessary to perform a human factors study so that an appropriate specification could be written for the appropriate hardware, software, and system requirements. The requirements for economical cost and the ability to use the system in the workplace resulted in specifying a portable compact hardware interface employing software compatible with PC's (personal computers). At the time this took place there was an extreme lack of PC-based empty-shell expert systems to serve this purpose. Developing a portable system with PC-based software and using it in the power plant workplace represented an important milestone in the use of expert systems.

The initial phase of this project resulted in a user interface which could be carried to the plant floor and plugged into a power and communications cable. this Phase I prototype system (Figure I) was tested at Jersey Central Power and Light (JCP&L) Company's Gilbert Station in Milford, NJ. The portable interface was used to interrogate the knowledge-base which resided on a host PC-computer in the control room.

The next phase incorporated all hardware and software into a single portable, brief-case size unit (Figure 2). This Phase II system had the advantages of:

1. Improved portability/mobility - all you need is a power connection

2. Faster response due to all hardware/software being self-contained.

The results of the field tests performed at JCP&L are shown in Table 1. The time required to trouble-shoot a ground fault is seen to be about the same for either the expert technician or novice technician, the reduced trouble-shooting time for the Phase II system also attests to its improved performance.

The User Interface

Although EPRI recognized the user interface to be an item crucial to the success of this project, it is gratifying to see the importance now being placed on user interfaces by others.

For example, Reference 1 cites the user interface to be of such importance that it can "make-or-break" an expert system:

> "THE USER INTERFACE IS CRUCIAL
>
> The user interface for an expert system is more than a display and an input device. Underneath the hardware is the software that makes the interface function for the application. It is the hardware and software together that determine the ease-of-use for the user. A poorly designed human interface will sink the expert system; it simply will not be used."
>
> R. S. Shirley

Reference 2 presents a compelling reason which could explain the difficulties encountered in moving expert systems from the laboratory environment into the everyday workplace:

Figure 1: Phase I Prototype Expert System Interface

Figure 2: Phase II Self-Contained Brief-Case Size Unit (SA•VANT)

"Failure to recognize the man/machine interface needs of the expert system users is probably the biggest reason for the disparity between the numerous expert systems which have been successfully developed in the laboratory and the small number which have actually made it into everyday field use. In the laboratory, expert systems tend to be used by people who love them and are tolerant of their idiosyncrasies. Outside the laboratory, they will only be used if people find them useful and easy to work with".

D.C. Berry and D.E. Broadbent

Industrial users, such as Alcoa Industries, also are appreciating the tremendous value of the user interface in terms of "getting the metal out the door". In Reference 3, Alcoa emphasizes that:

"Developing a meaningful interface is an important piece of the solution."

Peter Van Sickel

Applications and Future Expansions

Current applications have been for use in trouble-shooting gas turbine power plants (control system ground faults and turbine failure-to-start advisors).

Future expanded capabilities for this portable system include incorporating a data acquisition interface. Development of a vibration analysis expert system for gas turbines is planned for next year.

Other applications which can benefit from portability and interactive video may be installed as they are identified. Expert systems developed elsewhere have been installed and made operational in less than a two hour period.

References

1. Shirley, Richard S. Final Report on a Real-Time Diagnostic Expert System. Paper Presented at EPRI 1989 Power Plant Control Conference, Miami, Feb. '89.

2. Berry, D. C. and D. E. Broadbent. "Expert Systems and the Man/Machine Interface. Part Two: The User Interface." Expert Systems, Vol. 4, No. 1 Feb. 1987 pp. 18-28.

3. The Periodical for Symbolics Users, Vol. 3, No. 1, February 1989.

FIELD TEST OF GAS TURBINE EXPERT SYSTEM (GTES) AT JCP&L - GILBERT STATION

System Utilized	Average Time to Trouble-shoot Ground Fault: Expert Technician	Novice Technician
Man's own knowledge	60 min.	couldn't do
GTES - prototype	60 min.	65 min.
GTES - phase II	25 min.	26 min.

Table 1: Results of Field Test

EPRIGEMS™: Expert Systems for Technology Transfer

D. CAIN, E. CHOI, M. DIVAKARUNI, V. LONGO, T. WILSON, and B. BRAITHWAITE
Electric Power Research Institute
EPRIGEMS Project
3412 Hillview Avenue
Palo Alto, California 94304, USA

ABSTRACT

Expert systems are often viewed as an exotic technology, operating on specialized machines, involving expensive software, and requiring specially trained people. This paper suggests an alternative perspective. Expert system technology can be used in relatively sophisticated computer applications that run on personal computer (PC) installations. "Low tech" expert system technology can be successfully alloyed with more conventional computer programs, resulting in a hybrid concept for PC and workstation applications. The EPRIGEMS project at EPRI is developing this hybrid approach to package and transfer the results of R&D project as highly integrated, easy-to-use PC software. These software applications employ expert systems techniques to guide users in the solution of complex problems.

INTRODUCTION

During the past several years the notion of dedicated expert systems on specialized machines, embodying the knowledge of a single human expert, has been supplanted by hybrid system concepts. These systems combine expert systems and conventional computer technologies derived from a variety of sources. Hybrid expert systems embody knowledge, but not necessarily the knowledge of single human expert; they run on conventional computer hardware and interface with other programs and data streams, as well as the interacting with users. The EPRIGEMS project at EPRI is keying on hybrid expert systems as a means of configuring EPRI R&D technology and transferring it to utility users. EPRIGEMS symbolizes the extraction of valuable bits of information from EPRI research projects and cutting and polishing them into modules of compiled knowledge.

To apply EPRI research results in the past, utility engineers and planners usually read voluminous EPRI reports, consulted with EPRI project managers, and attended a seminar or two. Now, or in the near future, using expert systems as a guidance mechanism, they will be able to solve a problem, draw a conclusion, or implement EPRI technology right at their desks on personal computer (PC) systems. Interactive electronic handbooks, intelligent database access systems, integrated workstations, and computer-based instruction programs are examples of a new product line EPRIGEMS is developing.

This paper introduces the EPRIGEMS concept as a practical application of hybrid expert system technology, including the design philosophy that EPRI is using, the role of an intelligent session manager in interactively guiding users, software development environments, and example applications.

DESIGN PHILOSOPHY

In the utility industry, as well as in the engineering profession generally, getting others to apply complex technology reliability and efficiently is a major challenge. In contrast to the "classic" artificial intelligence problem of cloning knowledge resident in people's heads, the utility problem is often one of applying technology that already exists in a concrete form. This may be: a computer code or back-of-the-envelope calculation; small database or look-up table; graphic or characteristic curve; procedure or flowchart; text-based instructions or handbook. Very often, to solve a practical problem, one needs to apply some or all of these different resources, interactively.

In EPRIGEMS the approach has been to configure simple expert system(s), serving as navigators between "islands" of technology, rather than recasting existing technology into rules or other knowledge representations commonly used in expert systems. The results of EPRI projects are often manifested as analysis programs, text information, graphics, small databases, decision flow diagrams, or combinations thereof. These are the so-called technology islands. What is lacking is the means for navigating between them in order to achieve solutions to real problems. EPRIGEMS provides a framework for merging these technologies and orchestrating a solution to utility problems.

Each EPRIGEMS application is intended to be a compact, self-contained tool, known as an EPRIGEMS module. EPRIGEMS modules are designed to run on standard personal computer (PC) hardware, because utility personnel have these machines readily available to them and increasingly depend on them for day-to-day job functions. High-end workstations are rarely found in utility organizations. Artificial intelligence workstations are rarer still.

Current PC architectures impose significant limitations on expert system capabilities, both in terms of processing speed and memory management. However, this situation is somewhat ameliorated in EPRIGEMS by the fact that simple expert systems are used to link traditional programs and data structures. Moreover, with the introduction of new PCs and operating systems the performance gap between PCs and workstations is expected to shrink. The strategy in EPRIGEMS, then, is ride the crest of this technology wave, using applications design and software tools that run on PC's but which are upward compatible.

Given the task of providing intelligent problem solving tools that utility personnel can use on their PCs, a set of general design goals was developed for EPRIGEMS. These are shown in table 1.

Table 1: EPRIGEMS Design Goals

Standard "look and feel"	All EPRIGEMS Modules will have a similar appearance, not only to facilitate product recognition, but to give utility users assurance that, having successfully used one EPRIGEMS module, they can readily use any other module.
Upward Compatible	The EPRIGEMS designs will accommodate anticipated improvements and downstream computer technology innovations.
Intelligent Control	Principles of artificial intelligence will be used to create high-level problem-solving guidance; however, individual elements of a solution may be supported with traditional programming methods.
Development Flexibility	EPRIGEMS architecture will accommodate a variety of applications software and database types (as might develop from EPRI research and development) with capability to draw and use data and analysis results in problem solving.
Hybrid Capability	Developers will be able to use any software or software tools and tailor EPRIGEMS modules to specific applications, subject to minimum EPRIGEMS product specifications.
Output Capability	Where graphics output is available, a means of hard copy reproduction will be provided.

One of the important philosophical distinctions in EPRIGEMS, relative to common practices in the artificial intelligence community, has to do with the so-called "knowledge engineer". Whereas, large and complicated expert systems require specially trained AI personnel who understand the intricacies of knowledge extraction and representation, EPRIGEMS modules generally do not. Since EPRIGEMS modules feature fairly uncomplicated knowledge bases that link conventional programs and databases, it is well within the skills of traditional programmers and applications development engineers to master and apply the necessary expert system techniques. Considerable evidence from EPRI R&D projects developing expert systems applications seems to bear this assumption out.

EPRIGEMS SESSION MANAGER

The Session Manager is the nucleus of any problem-solving session in EPRIGEMS (see figure 1). It handles the communication between the user and various services, and inter-communications between services during a session. These services may include small expert systems, analysis programs, database retrieval, text handling, graphic displays, etc. The Session Manager exercises flow control, with means for storing and passing information, as well as assigning temporary control to services that perform particular tasks. In a sense the EPRIGEMS Session Manager is a "meta" operating system which provides "tactical support" to the user who is solving a complex problem.

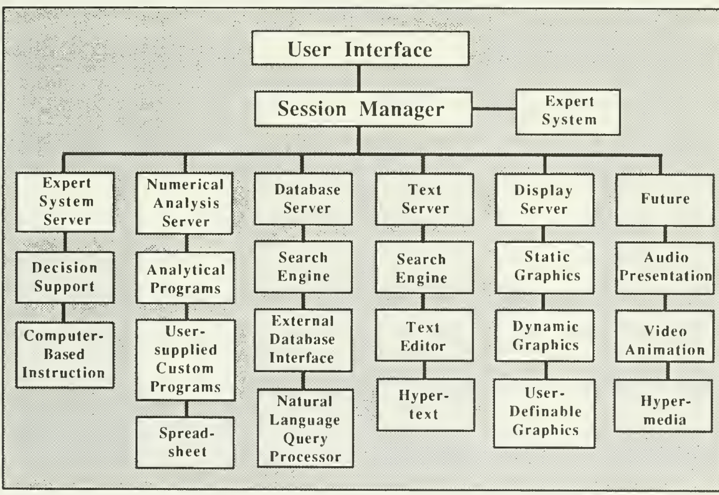


Figure 1: EPRIGEMS Architecture

At a superficial level, the Session Manager simply handles commands issued by the user via pull-down menu option selections, function-keys, form entries, etc. This capability allows direct user access to servers, as commonly allowed in any conventional software interface.

Complex problem-solving, however, does not always lend itself to this kind of "push button" operation. Complex problems follow irregular pathways, sometimes iterative or even recursive, that may opportunistically string together a variety of operations to arrive at a solution. This is illustrated in figure 2. In many traditional applications, these operations involve different software, requiring the user to pass results manually from one software application to another.

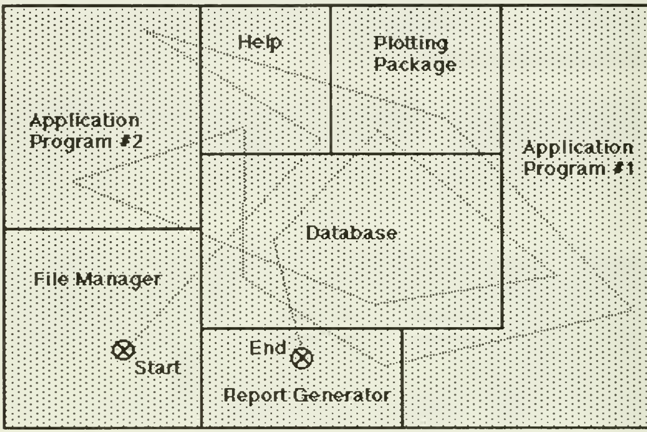


Figure 2: Software Solution Trajectory

Complex computer applications may require a virtuoso performance on the part of the user to achieve a satisfactory end-result. Novice and average users are left out; moreover, occasional users, once expert in using such software, cannot easily maintain their proficiency over the long term.

An "intelligent" Session Manager can alleviate this difficulty, at least in principle. This Session Manager not only handles direct user requests to initiate services, but also knows something about the nature of the problem being solved. It can monitor input, suggest alternative solution strategies, undertake a problem-solving session in an automated or semi-automated mode, understand the output, and present the output in a form that the user can digest. In its most advanced form, the Session Manager can "look over the shoulder" of the user and scale the level of support in proportion to the user's skill and complexity of the task at hand.

The following are hypothetical examples of intelligent session manager interactions with users:*

> "I noticed that a crack growth analysis has been recommended, based on an assessment of intergranular stress corrosion potential in your system. Would you like to do the crack growth analysis at this time?"

---

* The first person references in this examples are for illustrative purposes only. The use of the first person in human-computer transactions is highly controversial.

"Please fill in the next two forms and an input file for the XXXX code will be automatically generated. If you don't know the value that is appropriate for your plant, select "UNKNOWN." I will subsequently help you choose reasonable values, based on conservative estimates."

"The amount of radioactive iodine released appears to be in excess of the value implied by the plant technical specifications you supplied. Experience using this analysis program shows a significant reduction in the release if assumed feedwater temperature is increased. Do you want to try this?"

"In looking at your input so far, it appears that you have some expertise in soils analysis for transmission line applications. If you want, we can skip the following worksheets and proceed to the analysis itself. I will ask you for integral values as the analysis proceeds."

"We have been through a rather complicated analysis of underground cable systems design. Would you like me to recap the analysis path you used to show how your final design was achieved?"

Expert systems provide an excellent technical foundation for the intelligent Session Manager concept. Expert systems place a premium on highly interactive user-friendly interfaces, are capable of handling complex logic, support flexible data structures to accommodate input/output between the different servers, and provide excellent tracking and explanation facilities. Significantly, an array of sophisticated expert system shells are now available that greatly reduce the time and effort needed to build the kinds of intelligent support capabilities envisioned for EPRIGEMS Session Managers.

The role of expert systems in the Session Manager differs somewhat from the conventional notion of expert systems. To get the idea, one has to visualize a fairly broad, but not very deep knowledge base interfaced to the Session Manager block as shown in figure 1. This set of rules and objects does not actually solve the problem by inference, but interprets user commands and input values to organize and manage the overall solution process. By spawning a sequence of server tasks the actual solution is accomplished. The Session Manager's logical inference is continuous and may use output from a given server to redirect or opportunistically adopt a new solution scheme midstream. [Note that one server may be an expert system which, in the classical sense, may handle diagnosis, interpretation, etc. under the direction of an expert system Session Manager.]

Some of the most important expert system constructs used in Session Managers are the following:

- Object representation and message passing capabilities. Object representation is an alternative to rules for encoding knowledge. Objects possess attributes which can be interfaced to rules logic. In addition, objects may contain pointers to procedural code that can be triggered by a message from a rule associated with another object.

- Rule side effects. Rule side effects are one or more procedures, i.e., blocks of code that become active when that rule is satisfied during inferencing.

- Demon procedures. Demons, autonomous routines that are attached to object attributes, automatically activate when inferencing causes the attribute value to be accessed or the value itself is changed.

- External Interfaces. Built-in capability to query external databases or run external programs.

- Explanation. Facilities for expressing "why" a query is being made, or "how" a conclusion was reached.

Current Session Manager implementations use a well-integrated knowledge base architecture, exercising tight supervisory control over the solution process. An alternative architecture is a decentralized Session Manager, featuring a number of independent expert systems that are linked, via demon-like procedures, into the problem-solving scheme. A still more advanced Session Manager architecture would be a blackboard arrangement in which small expert systems, assigned to individual servers, cooperatively solve problems without need for a high level arbitrator.

The Session Managers developed for EPRIGEMS modules to date are fairly primitive, compared to capabilities outlined here. The evolution of the intelligent Session Manager concept will be an on-going EPRIGEMS development activity.

EPRIGEMS PRODUCT DETAIL

EPRIGEMS employs a standard "look and feel" interface [1]. The rationales for this are: product identification, ease of use, and economics. EPRI has produced a considerable number of PC-based software packages over the years. The lack of uniformity has engendered a "hodge-podge" image, due to the fact the every EPRI software package looks and works differently. Establishing a standard "look and feel" across a line of products addresses this problem, and also assures that a user who has applied one EPRIGEMS module can easily pick up and use another without having to master a new interface. Economic benefits derive from the fact that anywhere from 20-50% of the coding in PC software applications is related to user interface functions. EPRI R&D funding is being redundantly applied to interface developments by contractors who may be are more adept at research in a particular domain than designing good user interfaces.

The "look and feel" specification for EPRIGEMS reflects an industry trend towards window-based, pull-down menu interfaces. Although early EPRIGEMS modules were targeted for IBM-XT/AT machines running under DOS, there is a desire to maintain upward compatibility with Microsoft Windows/OS-2, as well as (possible) Macintosh applications of EPRIGEMS in the future. Accordingly, the standard top-level EPRIGEMS screen is as shown in figure 3a.

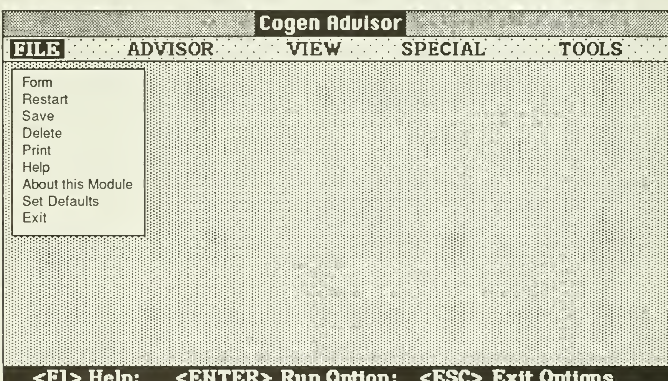


Figure 3a: Top-Level Screen and Pull-Down Menu

Using cursor keys (optionally a mouse) and the <ENTER> command, the user can select and initiate any menu option. The screens are spare in detail, and minimize the use of colors. A simple standard has been adopted.

In the top-level menu, the following conventions apply:

- FILE — Overall help, file management, and other housekeeping functions;
- ADVISOR — Analysis options and, in particular, expert problem-solving elements of the Session Manager;
- VIEW — Static information contained in the module, including text and data access, glossary information, and analysis results developed under ADVISOR;
- SPECIAL — Special purpose programs, including user supplied programs linked into the module using TOOLS;
- TOOLS — Utility functions used to support customization, configuration changes and special application programs installation.

The workspace below the main menu bar supports a variety of application-dependent features. Refer to figures 3b through 3d.

EPRIGEMS input conventions are intended to be as simple and fool proof as possible. User keyboard entries are automatically range and type checked; default values are provided. Multiple choice selection is employed for discrete values. Minimum keystroke design features facilitate ease-of-use and reduce typing errors. The escape key <ESC> exits any menu option or server. Function Key <F1> provides context sensitive help. In general, the use of function keys is minimized, avoiding the need for the user to memorize them or cluttering screens with their definitions.

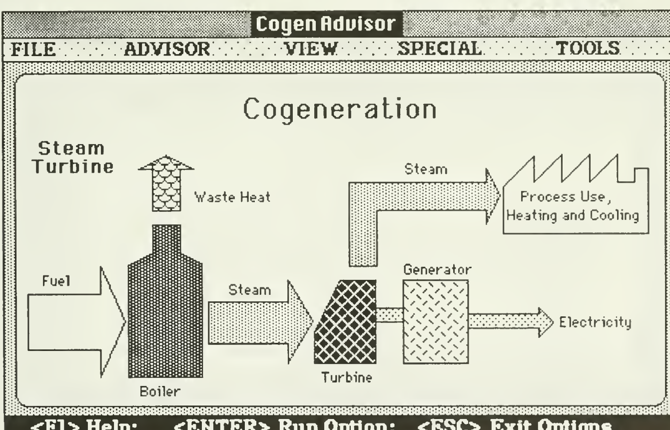


Figure 3b: Introductory Screen with a Color Graphic

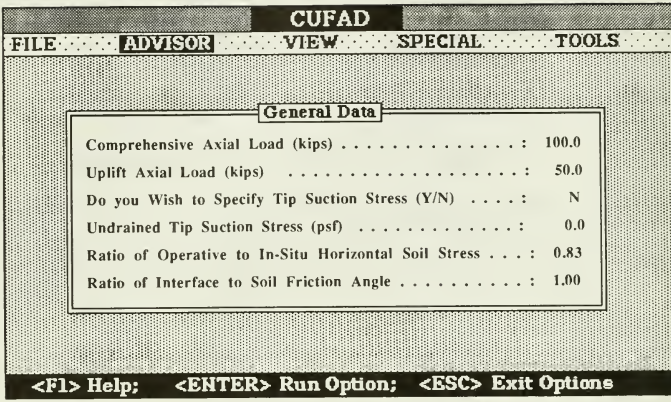


Figure 3c: Example User-Input Data Screen

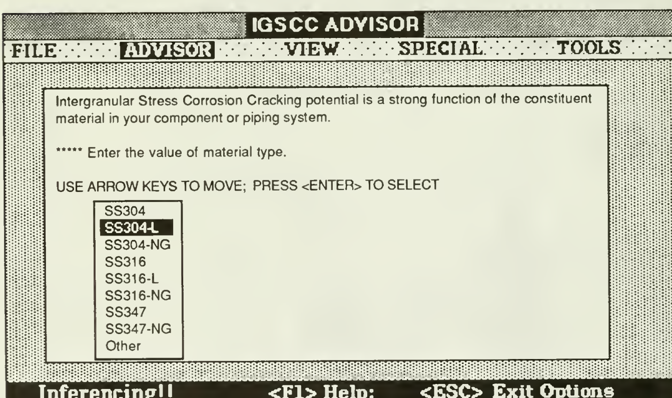


Figure 3d: Sample Screen for User Query Session

Each EPRIGEMS module is provided with an installation procedure, initiated by the command "EPRIGEMS". This procedure automatically unpacks files, creates a hard disk directory structure, transfers files from floppies to the hard disk, and transfers the user into the new directory structure. The module can be booted with a single command (usually keyed to the module name).

Packaging consists of a printed box, outfitted with slots for floppy disks and a pocket for user manual, reference card and supporting information.

SOFTWARE PLATFORMS

During the early phase of the EPRIGEMS project a concerted effort to evaluate commercially available PC software was undertaken. EPRIGEMS modules span a diverse set of potential applications; and, the software development skills of EPRI R&D contractors vary considerably. As expected, no single software platform was found to satisfy all of the prospective EPRIGEMS needs. Accordingly, an ensemble of software packages was ultimately identified and is being prelicensed for use in EPRIGEMS.

EPRIGEMS software in current use, or targeted for use, falls into four layered categories:

- Programming languages: Microsoft and Turbo "C"; Arity and Turbo Prolog; muLISP.
- Expert system shells: Nexpert/Object; SMART; PC Expert.
- Application development environments: Professional Applications Development Language (PADL Plus); EASE+.
- Miscellaneous: Graph-in-the-Box Analytic, Packarc, Dr Halo, etc.

In the base programming languages, symbolic processing capabilities and facilities to link with or interface to other software is critically important. Among these, "C" is considered the quintessential low level language due to its compactness, portability and power. Efforts are underway to establish a "C" library that fully supports the EPRIGEMS look and feel, and also includes a variety of utility functions for data handling, graphics and text management, etc. An off-the-shelf "C" toolkit will be acquired and upgraded for this purpose.

There are a plethora of good expert system shells for PC application. The three packages selected for use in EPRIGEMS range from relatively simple to sophisticated. Each shell is highly adaptable with sense that access to the underlying programming language or well-documented interfaces are provided. [It is important to note that Prolog is not only a programming language, but is also equivalent in many ways to expert system shells. It is regarded as such in EPRIGEMS.]

The application development environments provide high level facilities for constructing finished EPRIGEMS modules. They have been successfully used in past EPRI R&D projects to produce successful software products. However, prior applications have focussed primarily on interfacing analytic programs written in FORTRAN, etc. Work is underway to: (1) extend these products by interfacing with one or more expert systems shells used in EPRIGEMS; and (2) modify the user interface to comply with EPRIGEMS "look and feel" specifications.

A discussion of EPRIGEMS software would not be complete without touching on the gaps. At the present time no satisfactory package has been found that supports hypertext applications on IBM-PCs; yet, hypertext capability is a potentially powerful adjunct to the EPRIGEMS concept. Likewise, no general purpose package for intelligent text search and retrieval has been found, although some promising products are under investigation. EPRIGEMS has not yet found a stand-alone utility package designed for handling external queries to all (or many) of the popular PC databases. Finally, EPRIGEMS has plans to evaluate and eventually incorporate an authoring package for computer-based instruction into the existing software ensemble. A survey is planned, but has not yet been initiated.

The shaded blocks in figure I represent software capabilities that are currently not supported by EPRIGEMS. The process of identifying, qualifying and prelicensing this software will be an on-going EPRIGEMS activity.

EPRIGEMS APPLICATIONS EXPERIENCE

There are currently ten EPRIGEMS modules under development. One module, which is a small expert system, has been released [2]. Four others are essentially complete and undergoing beta testing. Examples of modules being developed are:

- Boiler Maintenance Workstation. Combines expert system failure diagnosis, analytical codes and database facilities to provide an integrated facility for boiler maintenance on a personal computer system.
- Chexpert. A computer package which will enable utility engineers to qualitatively assess erosion-corrosion effects in their plants and determine what EPRI analysis methods and codes should be used to deal with them.
- Foundation Soils Advisor. Expert system integrated with analytical procedures for providing a consistent, reliability-based evaluation of soil properties in transmission structure foundation design.
- Groundwater Quality Protection Advisor. Provides a highly integrated tool for evaluating and assessing groundwater quality, including analysis of leaching, monitoring and chemical testing of coal ash ponds.
- Starrs: a Code for Analyzing SGTR Events. This computer code, originally developed for mainframe analysis of pressurized water reactor steam generator tube rupture (SGTR) events, has been downsized for IBM-PC applications. An new, user-friendly interface has been provided with embedded expert system capability.

A backlog of approximately 30 additional EPRIGEMS applications have been identified by EPRI R&D staff.

So far, it is clear that developing EPRIGEMS modules is technically feasible and that technical staff "buy-in" to the concept is achievable. There are, however, some open questions:

- the extent real development cost savings will accrue from standardized, recyclable software;
- whether EPRI R&D contractors who actually build the modules can master the software technology or if a stable of qualified subcontractors needs be cultivated;
- what types of EPRIGEMS applications are "winners" and "losers" from a utility point of view;
- the overall percentage of EPRI R&D projects that are amenable to EPRIGEMS.

CONCLUSION

In EPRIGEMS expert systems are used in the Session Manager as a potentially powerful means of orchestrating solutions to utility problems in a user-friendly fashion. The user doesn't know, and probably will not care, that an expert system is working in the background as a guide in order to arrive at the problem solution. EPRIGEMS is an example of the idea that expert systems technology can, and perhaps ought to be, a means to an end rather than an end in itself.

As one looks forward to the arrival of some of the new and very powerful computer workstations under development, there will be a mismatch between the gross computing capability offered and the computing requirements of most utility engineering applications. Many industry observers believe that increasingly sophisticated "intelligent" interface software will eventually soak up this spare capacity.

EPRIGEMS anticipates these developments, albeit at a low level in order to be compatible with personal computer systems of today. Although much remains to be learned from experience derived from producing EPRIGEMS modules and interactions with users, the EPRIGEMS approach does suggest an interesting development pathway that utilities and other organizations might consider for their software products. Prospectively, some ideas engendered by EPRIGEMS may also translate into valid research topics within artificial intelligence and other computer science disciplines.

REFERENCES

1. "EPRIGEMS Product Specifications," Draft, EPRI, November 1988.

2. EPRI NP-6058, "LIFEX: Expert System for Providing Guidance on Material Degradation Mechanisms," November 1988.

Boiler Maintenance Workstation—An EPRIGEMS Application

G. P. SINGH and D. A. STEINKE
Karta Technology, Inc.
San Antonio, Texas, USA

J. SCHEIBEL and S. GEHL
Electric Power Research Institute
3412 Hillview Avenue
Palo Alto, California 94304, USA

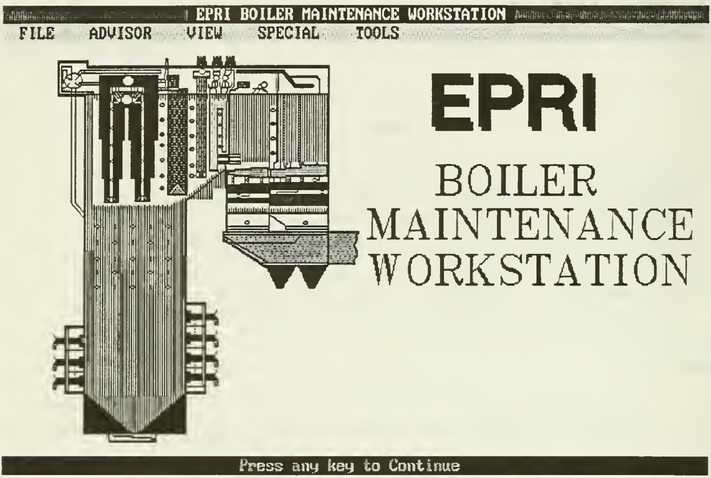
ABSTRACT

This paper briefly discusses EPRI's EPRIGEMS product specifications and the application of EPRIGEMS to the development of the Boiler Maintenance Workstation (BMW). The BMW, an EPRIGEMS product, operates on a personal computer and assists plant personnel in performing root-cause analysis, inspections, and repair decisions for boiler tubes. Its main purpose is to increase plant availability. This paper also discusses various modules incorporated in the BMW, and future plans for expanding the BMW.

INTRODUCTION

EPRI has developed a set of specifications to guide developers of software products intended for general utility applications. These specifications are referred to as EPRIGEMS. EPRIGEMS provides the framework for developing user-friendly software packages to deliver EPRI research and development project results. The goal of the EPRIGEMS specifications is to improve technology transfer.

An advanced application of these specifications is the EPRI Boiler Maintenance Workstation (BMW) (Figure 1). This EPRIGEMS product contains codes to address maintenance and engineering problems encountered in fossil-fired boilers. It is based on existing software for maintenance and life prediction and includes modules for tracking boiler-tube failures and repairs, analyzing ultrasonic thickness data from waterwall tubes, determining optimum inspection intervals based on economic analysis, and predicting remaining life of tubes exposed to high temperature creep. It also includes an expert system for determining boiler-tube failure mechanisms and aids plant personnel in conducting root-cause analysis.



<u>Figure 1</u> Opening screen of the EPRIGEMS Boiler Maintenance Workstation.

The BMW incorporates diverse user interfaces and presentation methods. The basic user interfaces are pull-down menus, pop-up menus, and data entry forms. A color spreadsheet-type interface is used for numeric and textual data entry and viewing. A graphic interface is also used to describe the different codes contained in the BMW. Other graphic data displays include bar charts, pie charts, and an isometric display of tube wall thicknesses. The BMW uses numerous fill-in-the-blank forms that allow the user to select information from a list of possible entries. These entries can be customized and/or expanded to meet individual plant requirements.

The primary goal of using graphic intensive displays and other user-friendly interfaces in EPRIGEMS products is to facilitate their acceptance by utility plant personnel. "Ease of use" is an essential requirement for plant maintenance codes. Maintenance personnel are responsible for a variety of activities and the use of specialized software occurs infrequently.

EPRIGEMS PRODUCT SPECIFICATIONS

EPRIGEMS specifications define a computer-based technology transfer mechanism to deliver EPRI research and development results to utility end-users. A few of the items described in the EPRIGEMS product specifications are:

- Problem Closure
- Standard "look and feel"
- Intelligent Control

An EPRIGEMS product should summarize research results that solve utility problems. Each module may combine information from various EPRI reports and analysis functions found in EPRI codes to address a particular utility concern. These modules can be updated as new technological advances are made.

All EPRIGEMS products will have a standard "look and feel". This not only provides product recognition, but more importantly, after becoming familiar with one module, utility users can readily learn another. Some of the major components of the EPRIGEMS "look and feel" are the use of pull-down menus, pop-up menus, forms, context sensitive help, graphics, and hypertext. The product specifications also define some of the standard features and options which should be present in most EPRIGEMS modules.

The intelligent control component refers to the use of an expert system to guide the user in determining a solution to a problem.

There are many levels at which this may be carried out. For example, an expert system could prompt the user for the type or area of the problem they wish to solve. Other problem-related information which could be acquired are: operation conditions, past history, and the amount and type of data currently available. Based on this information, the expert system would advise the user on the necessary steps in solving the problem. This could include a request for more data, suggestions on a sequence of codes to execute, and/or a list of applicable EPRI reports for reference. Once the suggested actions are performed, the expert system would use the results to make a determination.

BOILER MAINTENANCE WORKSTATION OVERVIEW

The major goal of the BMW is to provide solutions and aid in preventing, recording, and analyzing boiler tube failures using a user-friendly PC-based software system. The users of this system range from plant maintenance personnel to engineers and managers. The BMW platform is an AT or 386 IBM (or compatible) computer. An EGA monitor and graphics card are also required along with a printer for making hardcopies of data and/or to print reports. An HP Color PaintJet printer can be used to make copies of color graphic information.

The BMW integrates several previously developed codes which address boiler tube maintenance problems. The basic algorithms for the codes WW TUBE CONDITION, INSPECTION ECONOMICS, TUBE RECORDS, and TUBELIFE were developed under previous EPRI research projects while the expert system, ESCARTA, was acquired under a licensing agreement. The development considerations and a brief description of each of the BMW codes are discussed in the following sections.

Development Considerations

In developing the BMW the need to complete a user-friendly product in a limited time and within a fixed budget proved to be no easy task. A program's development time increases with its user-friendliness. Because of prohibitively large development costs and time, starting from scratch was not an option. Thus, finding the right tools to adapt existing software became extremely important. To conform to the EPRIGEMS standards, a very flexible user interface package was required. Fortunately one was found which provided the basic features. This "C" user interface library, "C-SCAPE" from Oakland Group, provided source code and after substantial modifications it was able to meet all of the EPRIGEMS user interface specifications. For developing a database, another "C" library, dBCIII from Lattice, was utilized. It provides dBASE III compatibility. Other graphics libraries were looked into, but the one included with the Microsoft "C" compiler proved to be appropriate for current needs.

Session Manager

The Session Manager provides information on each BMW module, overall help, a glossary of terms, and acts as a front end to the other EPRI codes included in the system. The user manipulates the cursor keys to highlight the code icon of interest and presses ENTER to display a brief synopsis of the program, i.e. why, when, and how to use the module. The selection screen for the Session Manager is illustrated in Figure 2. An example screen for one of the modules is shown in Figure 3. The menu in the upper right-hand of the screen allows the operator to select more detailed information on the module.

Tube Records

Tube Records is a database for tracking and recording tube failures, repairs, and analysis information. The information stored includes tube location, failure date, failure mechanism, root cause, man-hours for repair, and power lost in a forced outage. The database also tracks boiler tube repairs and associated information such as repair/replacement date, location, tube specifications, repair method, cause of repair/replacement, date of repair/replacement, and life of previous tube. It also is capable of recording analysis information such as analysis date, boiler location from which a sample was taken, results of metallurgical analysis, etc.

The database is designed to minimize the amount of typing and manual data entry by using pop-up selection lists for fields which have a known set of values as shown in Figure 4. This greatly improves data integrity by reducing the possibility for error, and makes data entry easier. If the values found in the selection lists are not adequate, users may add necessary options which will be displayed whenever the selection list is called. The user can also customize the database by adding fields to the basic version.

The database has standard functions such as: search, sort, sum, average, and count. Records may be viewed and printed singularly as a form or in a tabular format. Reports can also be generated with bar and pie charts.

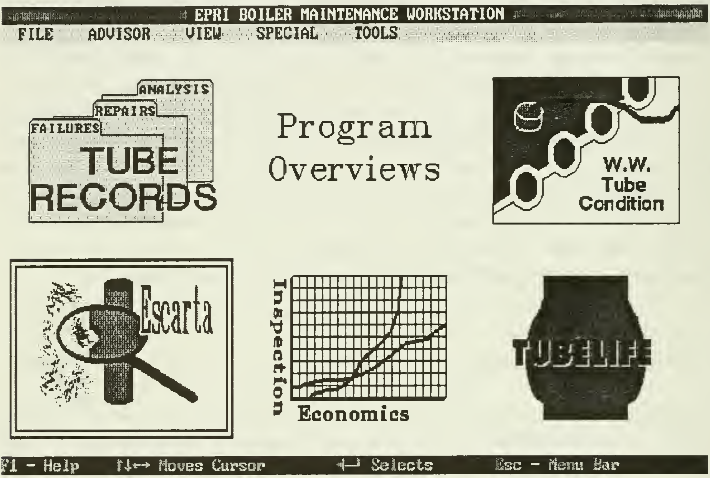


Figure 2 Session Manager Graphical Selection Screen for Program Overviews. This depicts one of the graphical interfaces used in the BMW.

WW TUBE CONDITION

WW TUBE CONDITION is used to help plant personnel analyze ultrasonic tube thickness data in the boiler waterwall and plan future boiler-tube inspections, maintenance, and tube replacements. Some of the functions of WW TUBE CONDITION are:

> Store tube thickness data obtained from ultrasonic examinations. Examination data may be entered

automatically via a file import mechanism or entered manually from a built-in, spreadsheet type interface.

- Calculate tube wastage rate from two examination data sets.

- Calculate the wastage rate of a specific area of the waterwall.

- Calculate remaining life or future thickness based on the calculated wastage rate.

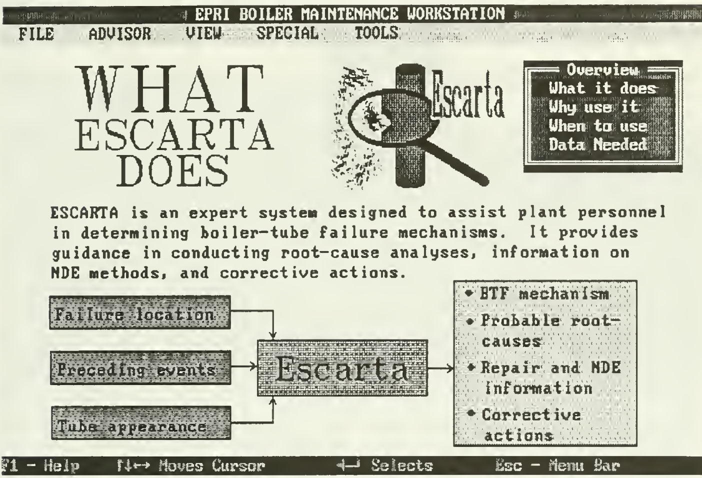


Figure 3 ESCARTA Program Overview "WHAT" screen. Information on what a code does, why use it, when to use it, and what data is needed can be displayed.

- Display thickness and remaining life information of the waterwall in three formats: graphically, isometrically, or as a spreadsheet. The data is displayed in multiple colors that correspond to different thickness thresholds to readily allow the identification of trouble spots.

- View and edit thickness data in the spreadsheet interface. Textual information may be attached to examination locations for record keeping purposes as shown in Figure 5.

Users can switch between the graphics and spreadsheet displays and can select different data sets. This facilitates quick comparisons of data such as current and calculated future thickness or previous and current thickness.

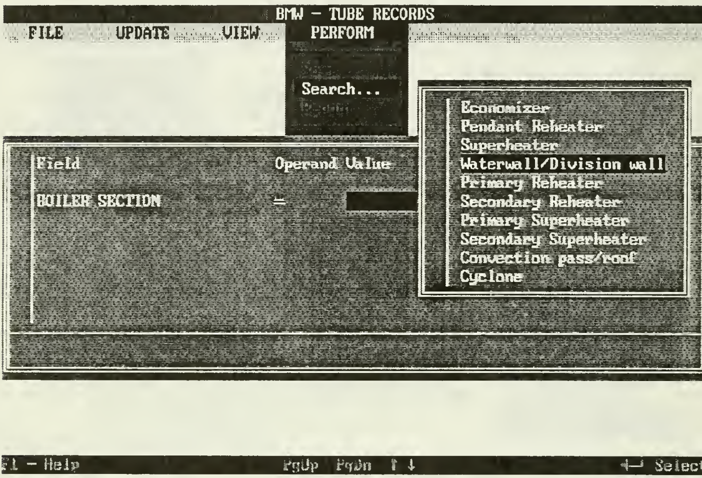


Figure 4 TUBE RECORDS Pop-Up Selection Menu. This provides easy data entry and also enhanced data integrity. Selection lists may be user customized as needed.

ESCARTA

ESCARTA is an expert system designed to help maintenance personnel analyze boiler-tube failures (BTF). ESCARTA is based on the knowledge compiled in EPRI Report CS-3945 Manual for Investigation and Correction of Boiler Tube Failures. It emulates the capabilities of human experts in BTF analysis. ESCARTA can be used to quickly determine the tube failure mechanism, provide preliminary leads for root-cause analysis, and recommend verification and corrective actions including NDE methods and repair procedures.

ESCARTA can be used by power-plant generation and operations managers, maintenance staff, and other plant personnel who are not experts in BTF analysis. ESCARTA determines failure mechanisms based on tube failure location, appearance of the failed tube, and events preceding the tube failure. Diagnosis is conducted by obtaining information using IF-THEN rules. ESCARTA determines one of 22 possible failure mechanisms and recommends a course of action.

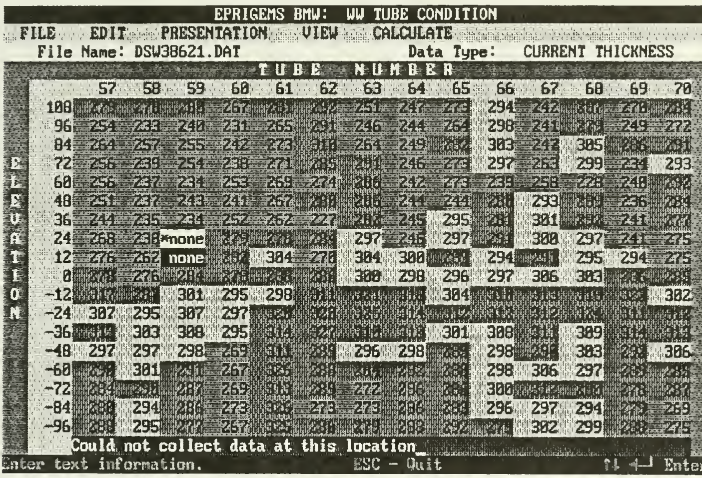


Figure 5 The WW TUBE CONDITION spreadsheet interface depicts the entry of textual information which is indicated on the screen with a preceding asterisk.

The rule base is divided into four distinct sections: waterwall, economizer, superheater, and reheater. Specific failure location questions are asked. For example, locations in the waterwall are referenced relative to the burner level, in straight runs, bends, welds, welded attachments, etc. Once the exact location of the failure is known, questions about events leading to the failure are asked. These include questions about such events as a drop in water level, flame impingement, and high heat flux area. It should be mentioned that in many instances it is not possible to confirm the existence of certain events. ESCARTA has been designed to operate under such uncertainties. Once the failure location and events have been ascertained, emphasis is placed on information about the appearance of the failed tube. An optional random access slide projector/viewer is available which reinforces the appearance descriptions with high-resolution slides of various failed tubes.

After the failure mechanism is determined, context sensitive information can be accessed. Examples include the root cause(s) of the failure, nondestructive evaluation methods, metallurgical tests, repair procedures, references, and corrective actions (Figure 6). Users can access context sensitive information for various failure mechanisms at any time. By making this information readily available, ESCARTA makes an excellent training tool for teaching maintenance personnel and others about the cause and effect relationships that are used in analyzing tube failure mechanisms and in conducting a root-cause analysis of tube failures.

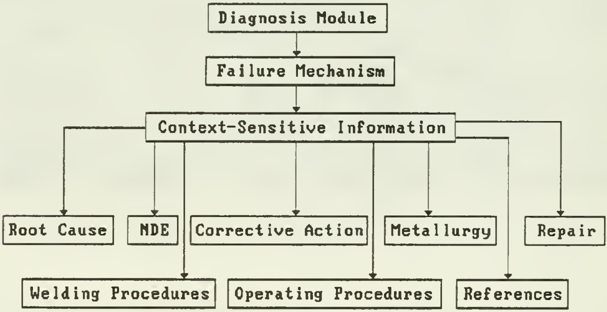


Figure 6 ESCARTA Structure and Function. ESCARTA provides context sensitive information which can be customized to include detailed company procedures.

Inspection Economics

The Inspection Economics module optimizes the length of the interval between boiler thickness examinations to provide the greatest economic benefit. It bases its calculations on examination costs, repair costs, and failure costs. The tube wall thickness distributions and wastage rate are also needed. This information can be entered manually or imported from data files produced by the WW Tube Condition code.

Monte Carlo simulation is used to determine the optimal examination intervals. The tube thickness distribution(s) are graphically shown as the simulation is performed. Yearly costs for examinations, repairs, and failures are also displayed graphically.

The code is designed to allow a one-time entry of most of the pertinent information. This information can be saved and recalled at will. Once the default information has been entered, changing just a few parameters will allow "what if" calculations to be performed rapidly.

TUBELIFE

The TUBELIFE module determines the remaining creep life of ASME SA213-T22 superheater or reheater tubes which have had significant service exposure. The methodology on which this is based is found in EPRI Report CS-5564, Remaining Life Assessment of Superheater and Reheater Tubes.

The remaining creep life is calculated from hoop stress and temperature histories. Hoop stress is determined from tube wall thickness measurements, while the temperature is estimated from the thickness of the insulating steamside oxide scale.

FUTURE PLANS

A utility users group is being organized to validate the current BMW modules. Each utility has its own operating and maintenance procedures and availability goals factored in to the workstation. Applications range from plant installations for quick response to routine maintenance to centralized engineering installations for monitoring all boilers within a generation system. Such diverse requirements along with various boiler design features will fully test the BMW.

Expected areas of new code development include the analysis of thick-walled component damage (headers, drums, steamlines), boiler performance, and a maintenance advisor to assist personnel in planning and executing maintenance programs and procedures. Further, developments will include a graphics database to show tube failure, repair, and remaining life information. The graphics would be customized for each boiler. A training module is also planned to assist plant personnel in using the BMW for problems specific to their plant.

CONCLUSION

In the past, as the complexity of the problems solved by computers increased the difficulty of using the computer codes also increased. To counter this, EPRI has developed a guideline or set of specifications named EPRIGEMS. The EPRIGEMS product specifications define an easy-to-use, computer-based technology transfer vehicle to deliver EPRI research and development results. EPRIGEMS combines standardized user interfaces, graphical interfaces and displays, expert system technology, extensive on-line help, and analysis codes to solve specific utility problems.

The EPRI Boiler Maintenance Workstation specifically addresses problems in fossil fired utility boilers. The BMW includes a database for tracking boiler tube failures and repairs, and codes for analyzing ultrasonic thickness data from waterwall tubes, determining optimum inspection intervals based on economic analysis, and predicting remaining life of tubes exposed to high temperature creep. It also includes an expert system for determining boiler-tube failure mechanisms and aids plant personnel in conducting root-cause analyses.

Future goals include the addition of thick-wall analysis codes, performance monitoring codes, an expert system based "maintenance advisor", a training module, and a graphically driven tube failure, repair, and remaining life database.

REFERENCES

1. "EPRIGEMS Product Specifications - Draft 2.0", Electric Power Research Institute, Palo Alto, CA. November 1988.

2. J. J. Hanley, A. S. Birring. "Fossil-Fired Boiler Tube Inspection: Volume 2: Nondestructive Testing Computer Programs", Southwest Research Institute, San Antonio, TX, July 1988.

3. G. P. Singh, S. M. Gehl, J. R. Scheibel. "A User-Friendly Expert System for Diagnosing Boiler-Tube Failures in Fossil Power Plants", Seminar: Expert Systems Applications in Power Plants Notebook, EPRI, Palo Alto, CA. May 27, 1987

4. S. R. Paterson, S. Schipper. "TUBELIFE: Remaining Creep Life Assessment Code for ASME SA213-T22 Superheater and Reheater Tubing - Programmer and Users Manual", Aptech Engineering Services, Inc, Sunnyvale, CA. June 1988.

Evaluation of an Emergency Operating Procedure Tracking Expert System by Control Room Operators

J. F. CHENG, RAY CHIANG, and C. C. YAO
Taiwan Power Company
242 Roosevelt Road
3rd Section
Taipei, Taiwan, ROC

ANTHONY J. SPURGIN and D. D. ORVIS
Accident Prevention Group
11545 West Bernardo Court
Suite 100-4A
San Diego, California 92127, USA

BILL K. H. SUN and DAVID G. CAIN
Electric Power Research Institute
3412 Hillview Avenue
Palo Alto, California 94304, USA

C. CHRISTENSEN
General Electric Company
175 Curtner Avenue
San Jose, California 95157, USA

ABSTRACT

Taiwan Power Company has conducted an extensive program at the Kuosheng Boiling Water Reactor Simulator facility to install and evaluate the EPRI-developed Emergency Operating Procedures Tracking System (EOPTS). The EOPTS is a real-time expert system that assists reactor operators in monitoring and carrying out EOPs during reactor transient events and accidents. The evaluations, which used human factors technology, were performed for six accident scenarios, with operator crews divided into two groups, one using EOP flow charts directly and the other using the EOPTS. Results show that use of the EOPTS can reduce the rate of errors as well as the time required for operator responses. This evaluation indicates that the EOPTS meets its design goals of enhancing the operator responses to accidents and in doing so significantly increases the reliability and safety of plant operations.

BACKGROUND

EMERGENCY OPERATING PROCEDURES TRACKING SYSTEM

Nuclear plant safety systems include automatic protection systems and trained operators who follow approved emergency operating procedures (EOPs). For complicated transients requiring operator intervention, effective use of EOPs is a crucial part of the emergency response process. Because EOPs can be rather complex, selecting the correct procedures and applying the associated decision logic impose considerable operator burden. Inevitably, this effort takes time that could be better spent employing measures to control and stabilize the plant.

Using expert system technology, a means is developed to interpret and compile emergency procedure logic into a compact, fast-running software module that interfaces with and uses the same database as the safety parameter display system (SPDS). As programmed, the system allows multiple user access - for example, in control rooms and technical support centers. It provides real-time notification of emergency procedure steps, on-line explanations of messages, priority filtering, and checking of data quality.

The EOP tracking system (EOPTS) is based on the emergency procedures guidelines of the BWR Owners Group, using the EOPs of the Taiwan Power Company's (TPC) KuoSheng Boiling Water Reactor as a specific model. (1, 2, 3, 4) The system provides an on-line display of the appropriate steps in these EOPs, traversing the entire procedures logic at short time intervals. By enhancing operators' abilities to interpret and apply these procedures, the computer-based tracking system developed by EPRI can help reduce human error.

TEST DESCRIPTION

Initial EOPTS evaluation tests were conducted at the Taiwan Power Company's KuoSheng simulator facility in September, 1988. The tests were performed with three of the crews of the two-unit Kuosheng BWR/6 plant. For the tests, each full crew was split into two four-member crews designated "A" and "B", making six test crews in all. Each crew thus consisted of two control operators and two supervisors (at least one Senior Reactor Operator).

The second series of tests was conducted at KuoSheng in February, 1989. The tests were performed with six shifts and each shift was divided into two four-member crews also, for a total of twelve test crews.

For the first series of tests, one of the A or B crews would use the EOPTS and the other crew would use the Flow Chart. Crews using the EOPTS were instructed to follow the messages verbatim. Each of the six crews was exposed to two scenarios labeled as Scenario 3 and 4. Two crews were also exposed to scenarios 1 and 2. The four scenarios are:

1. Anticipated Transient Without Scram (ATWS)
2. Radiation Release Accident Due to Steamline Break
3. Loss of Emergency Core Cooling System (ECCS)
4. Loss of Reactor Pressure Vessel (RPV) Level Indication

It is important to note that none of the crews had any substantive prior practice using either Flow Chart or EOPTS.

Subsequently, it was decided as a result of these initial tests to do two things; 1) increase the degree of training of the crews in the use of the EOPs using flow charts, and 2) to expose the crews after this increased training to two difficult sequences. During this second series of experiments, crews would be observed using either flow charts or the tracking system. Analyses of the experiments are given in this paper.

For the second set of experiments carried out in February 1989, two new scenarios were designed. These are:

5. LOCA with drywell/primary containment hydrogen control
6. ATWS with abnormal suppression pool level

Again, for the second series of tests, one of the A or B crews would use the EOPTS and the other crew would use the Flow Chart, with each of the twelve crews exposed to scenarios 5 and 6. Crews were given additional training (one-two months) in the use of the EOPs in flow chart form prior to the second test series, as per a request to TPC from the Republic of China Atomic Energy Commission.

DATA COLLECTION

Two measures for evaluating EOPTS effect on crew performance were established during test planning:

1. Number of deviations from the EOPs, and
2. Time responses of the crews in applying EOPs to diagnose and perform appropriate control actions

Data on EOP deviations were obtained directly from printouts of the EOPTS message recording feature. Messages appear as "NEW" entries when conditions call for them and appear with "DEL" prefix when the action has been completed or conditions change. Reconciling the "DEL" vs. "NEW" message pairs in a printout shows which messages remain active in the EOPTS at the time the scenario is terminated by the simulator instructor. The EOPTS was operating during all scenario runs even when the crew was using the Flow Chart; hence, this EOPTS message reconciliation was made for all runs. (The EOPTS printout also provides times when the NEW and DEL messages occur which is used to supplement other timing data.) Data on EOP deviations was supplemented with data obtained during the debriefing interviews of the crews.

The primary means for obtaining timing data was human observers. Several of the authors and members of the TPC team recorded times of cues and crew actions on forms prepared for each scenario. Stop watches were used to note the elapsed time from the start of the scenario (or time of reactor scram) to each prescribed cue and action. The data were analyzed subsequently to compare the time intervals between selected cues and actions for crews using the EOPTs and Flow Charts, respectively.

Other data included observation of Human Factors information using a prepared form and crew experience/background statistics.

As a result of the initial experiments, a new form was developed which has as its objective the need to determine the likely cause of crew deviations from procedures, and if the crews recovered from these deviations. This "Error Type-Cause Matrix", or "Slip Matrix", was completed by the observer during each experimental run. The root cause analysis was carried out by the observers following each test scenario.

This data is useful in determining the efficacy of the EOPTS versus the EOP Flow Charts.

RESULTS

Results are reported for both measures of EOPTS evaluation: number of deviations and comparison of response time data. Since the initial test series provided only one to three data points for each test, the statistical basis is weak. Nevertheless, the preliminary results indicate a performance improvement for crews using the EOPTS.

In addition to the results from the initial experiments, some results from the later experiments are given; here the statistics are better since there are 12 crew data points per scenario. The complete analyses for these scenarios have not been completed, but some early results are given below.

TIME COMPARISONS

To compare the EOPTS against the Flowcharts, a time difference for a cue-action pair (human interaction) was used. The time difference is the time between the cue and the operators' taking an action. Within the time interval the operators need to recognize the cue, find the appropriate steps in the EOPs read them, and execute the action. One cue-action pair (human interaction) which spans the use of an EOP segment was selected for each scenario. Results for scenarios 3, 4, 5 and 6 follow below.

For Scenario 3, the human interactions cue is "water level reaches top of active fuel" and the action is "initiate emergency depressurization." The analyzed time data are shown as follows:

Scenario 3	Number of Crews	Tavg+ Sec	SD* Sec	Ratio SD/Tavg
Using EOPTS	3	194	77	0.4
Using Flow Chart	3	465	475	1.0

+ Tavg = Mean of time interval between cue and action for n crews
* SD = Standard deviation of time interval between cue and action

The results indicate the average crew response time using the flow chart is about 2.5 times longer than for crew using the EOPTS. Further, the ratio of standard deviation to mean response time (normalized measure of variability) can be interpreted in the Human Cognitive Response framework to indicate a "skill" or "rule-based" type of cognitive behavior using the EOPTS (ratio of 0.4) while the crews using the flow chart indicate more "knowledge-based" (ratio of 1.0). (5, 6) Since the mean and SD represent only three data points, the statistical limitations must be recognized in reporting these results.

For Scenario 4 the human interactions cue is "reactor scram" and the action is "initiate emergency depressurization" after the dry well temperature exceeds the saturation temperature of the RPV. Results are similar to those reported for Scenario 3.

Scenario 4	Number of Crews	Tavg Sec	SD Sec	Ratio SD/Tavg
Using EOPTS	3	196	63	0.3
Using Flow Chart	3	770	659	0.9

For Scenario 5 a the human interactions cue is "Rx level drops below top of active fuel " or "drywell hydrogen level equals or exceeds the deflagration pressure limit." The action is "emergency depressurization".

Scenario 5	Number of Crews	Tavg Sec	SD Sec	Ratio SD/Tavg
Using EOPTS	6	82	48.8	0.6
Using Flow Chart	6	262.5	187.50	0.77

The results indicate the average crew response time using the flow chart is about 3.2 times longer than for crew using the EOPTS, a significant margin. The ratio of standard deviation to mean response time does not indicate a substantial difference between the EOPTS and Flow Chart crews, however those using the EOPTS do perform at a higher level of effectiveness.

For Scenario 6 the human interactions cue is "MSIV isolation/Rx scram" and the action is "Trip recirculation pump B."

Scenario 6	Number of Crews	Tavg Sec	SD Sec	Ratio SD/Tavg
Using EOPTS	6	94.2	47.3	0.51
Using Flow Chart	6	92.5	99.81	0.08

The results indicate the average crew response time using the flow chart is about the same as for crew using the EOPTS. The standard deviation indicates greater consistency amongst crews using the EOPTS. The ratio of standard deviation to mean response time does indicate a substantial difference between the EOPTS and Flow Chart crews within the Human Cognitive Response framework. Crews using the EOPTS exhibit a "skill" or "rule-based" type of cognitive behavior, while the crews using the flow chart indicate more "knowledge-based".

While not included herein for brevity, time results for Scenario 1 indicate similar improvements using the EOPTS. The results for Scenario 2 show essentially no quantitative improvement with the EOPTS; this scenario was relatively slow moving and not complex--essentially only a small portion of the EOPs had to be followed.

A few additional observations from Scenario 5 are worth noting. One critical measurement for this transient (LOCA with drywell/primary containment hydrogen control) is the concentration of hydrogen in the drywell (with the consequent risk of combustion). For crews using the EOPTS the maximum drywell hydrogen concentration averaged 5.9% (range 5.1% to 7.3%). For crews using the Flowcharts the average was 8.8% (range 7.2% to 10.0%). Moreover the latter data probably underestimates the actual concentration levels; values for three of the crews reached there maximum or were still increasing at the end of the parameter printout, and one crew "pegged-out" at ten (the parameter printout gave no values over 10%). This indicates a substantial risk from excess concentration of hydrogen in the drywell for crews using the flowcharts. Figure 1a and 1b give an example of this for two crews.

The difference in hydrogen drywell concentration in part may be attributable to the Tracking System's auto-monitoring of hydrogen levels, information immediately accessible by crews using the EOPTS. Crews using the flowcharts had to rely on a "back panel" hydrogen meter; observer comments indicate that several crews took time to locate it.

In Scenario 5 cumulative time below Top of Active Fuel for operators using the EOPTS was consistently lower than that for those using the Flowcharts (average of 92.5 seconds vs. 325 seconds; ratio of 1:3.5). This could be a significant factor in avoiding core damage during accidents. For this scenario minimum RPV level also did not fall as much for EOPTS crews than for those using flowcharts (-628cm vs -776cm). Moreover the readings for three of six crews using the flowcharts "pegged-out", meaning they exceeded the capability of the simulator to accurately represent the level beyond this value. This occurred with only one of the EOPTS crews. Figures 2a and 2b graphically depict this difference for two crews. (Note in Figure 2a (EOPTS crew) the RPV pegged out.) The data also indicates that crews using the EOPTS return to an original condition (recovery) faster than those using the EOPs in flow chart form.

DEVIATIONS FROM EOPS

Using the EOPTS' message status as a reference of performance, deviations from the EOPs were observed on the basis of unresolved EOPTS messages left at the end of the session.

At the conclusion of the scenarios for crews using the EOPTS, the EOPTS screen generally showed only EOP "entry conditions" as still being active, i.e., messages such as Entry to RPV Level Control, etc. For Scenario 3, one of the EOPTS crews had some additional messages remained that would have been resolved if the simulation were continued; these included messages like "put RHR in shutdown mode". Another crew had an unanswered "Ask User" message on the screen.

By contrast, all crews using flow charts had several unresolved messages on the EOPTS screen (monitored by the observer) at the end of both scenarios. For example, in Scenario 3 one crew had the message "Start D/G (Diesel Generator) II"; had "Initiate ADS (Automatic Depressurization System), " "Augment Depressurization", and "Put Mode Switch in S/D (Shutdown)"

For Scenario 4, all crews using the flow charts, "Stop CGCS (Combustionable Gas Control System)" remaining while none of the crews using the EOPTS had this message unresolved and two EOPs crews had the message "Trip Recirculation Pumps".

It is noted that for Scenario 2 involving the Radiation Release portion of the EOPs, experienced by only two crews, there was no difference in messages remaining for the crew using the EOPTS and the crew using the Flow Chart. This was explained by the crews who noted that (1) this portion of the Flow Chart is easy to follow because it does not involve simultaneous control/monitoring of RPV level, primary containment, etc., and (2) the transient was relatively slow.

Data on EOPTS message status for Scenarios 5 and 6 are still being reviewed.

Because the course of a transient and the appropriate EOPs may change from crew to crew depending on what and when crews do certain things, the messages remaining in the EOPTS may not all represent deviations relative to current conditions. But, if following the EOPTS verbatim is regarded as the standard of performance, then use of the Flow Charts leads to more deviations by crews. In the case of the first four scenarios this may be explained by the crews having had little prior practice with the EOPs. More recent experiments enumerated deviations from observer data as described in Section 2.3.

ERROR ANALYSIS

The second set of experiments enable an analysis to be made of the types of errors made by the crews in responding to the accident. These errors, such as failure to take the appropriate EOP step or missing a step, are recorded along with data on whether or not the crews recovered from their errors. Data was collected for 14 crew scenarios with and without the use of the EOPTS by the crews. The results are:

Total number of errors:
with Flow Chart: 23
with EOPTS: 11

Number of unrecovered errors (within time limits):

with Flow Chart: 15
with EOPTS: 3

It was also noted that the error tendency with flow chart use was different to that with EOPTS use. The majority of errors with the flow charts are procedural, whereas those with the EOPTS are mainly communication difficulties between crew members or errors of execution (slips) which are easily recovered.

QUALITATIVE OBSERVATIONS AND CREW COMMENTS ON EOPTS

Overall, crews using the EOPTS were able to use it successfully. Figure 3 shows the test setup at TPC's KuoSheng BWR simulator site, with human-factor observers in place, operation crews standby, and transient about to start. There were a few problems in use as noted by the observers and crews.

There were occasional problems using the MORE, WHY and ASK USER functions, especially during the more rapid transients. These problems were due to a combination of (1) lack of prior crew practice with the EOPTS and (2) design of the user interface which requires a somewhat confusing use of "function" keys on the keyboard. A simpler keyboard having only a few necessary keys labeled "yes", "no", "more", etc. would help.

The use of a relatively small CRT placed on a desk constrained the SROs from being more aware of the overall plant condition. Following instruction to use the EOPTS verbatim, the SROs tended to remain seated and use the EOPTS and RO feedback as their principal means of following the transient. Crews suggested that placing a larger CRT higher on the control board would allow them more freedom as well as permitting the ROs to see the EOP messages.

One crew noted that the design of the message hierarchy could be improved, particularly with respect to CAUTIONS. They could not easily relate a specific caution on the screen to a given action message; they suggested that the CAUTIONS be coupled with the message on the screen and not "piled up" with other cautions at the end of entry/action messages.

A cursory examination of the Observer Forms for each test indicates that crews using the Flowcharts exhibited a higher frequency of problems, confusion, or stress than did those using the EOPTS. The difference approaches a ratio of 3:1 for scenarios 5 and 6.

Several other parameters associated with the functioning of the EOPTS may be seen as impacting on crew performance. In Scenario 6 crews using the EOPTS resorted to SBLC (boron injection) at a higher rate than did those using the Flowcharts (two of six versus one of six). This is partially understandable as four of the six flowchart crews never reached a SBLC condition. However, this may also be attributable to the instructions given the EOPTS crews to follow it verbatim; hence when the request for SBLC appeared they responded immediately. A third EOPTS crew received the command "Initiate SBLC", but the conditions were borderline and the crew decided not to follow the command. A few minutes later the command to "initiate SBLC" disappeared. Crews using the flowcharts in similar circumstances may have been able to use some discretion in implementing SBLC, allowing the plant to retreat from SBLC conditions before they felt compelled to take action.

CONCLUSIONS AND RECOMMENDATIONS

The results of the limited set of tests indicate that use of the EOPTS improves crew performance in controlling complex accident scenarios in comparison to crews using Flow Chart EOPs. Although the statistical base of the initial transients is limited, preliminary comparisons of mean values and dispersion of crew response times in the Human Cognitive Response framework indicate that crews using the EOPTS (without much prior practice) operate in the "skill-" or "rule-based" cognitive domain as shown in Figure 4 (which should be expected when directed by an "expert system"). Crews using the Flow Charts, both with and without much prior practice, operate more in the "knowledge-based" mode, as shown in Figure 5.

The smaller standard deviations for crews using the EOPTS also demonstrates a greater consistency amongst this group. For the human interaction in Scenario 6 (trip recirculation pump B), although the crews using the flowcharts actually had a faster mean response time, the comparatively larger standard deviation indicates the existence of large outlier values and hence crew performance is likely to be less dependable.

The ability of the EOPTS crews to minimize drywell hydrogen concentrations in Scenario 5 may in part be attributable to the Tracking system's ability to auto-monitor such parameters and display them directly to the crew on a recurring basis, thus liberating the crew from the requirement of physically locating the appropriate meter, and reading and recording the data. This advantage should not be underestimated, and may in fact be a significant strength of the system. In complex and stressful accident sequences, reference to back panel data will be constricted by time limits and constraints on operator cognition from data-overload (as was apparently the case for those crews using flowcharts in Scenario 5). The Tracking System has the potential of averting this problem.

It should be pointed out that the data indicates that Human Interactions of relatively short duration (small time interval between the cue and action) generally favor crews using the flowcharts. This was particularly apparent in the results from Scenario 2 (Radiation Release). This may in large measure be accounted for by the fact that the Tracking System has a built in 15-30 second time-lag between the occurrence of an event and the systems ability to report it (due to the fact the EOPTS shares the computer with the Simulator, which takes precedence in task execution). Consequently, Human Interactions requiring a short time period are biased towards the flowchart operators, except in those cases where Tracking System crews "jumped the gun", and initiated an action prior to instruction from the EOPTS (the mode switch action in Scenarios 5 and 6, for example).
The results from the second series of tests corroborate the general conclusions from the earlier tests. The overall error rate with the EOPTS is significantly lower than with EOP flow charts. Of special note is the fact that the recovery rate is much higher in the case of EOPTS use, i.e. 4:1 versus 2:1.

Based on the results of experimental testing, the conclusion drawn is that the EOPTS has a marked effect on the performance of control-room crews. In general, crews using the device display greater consistency, have fewer discrepancies, and are more successful in recovering from discrepancies that do occur. This means that simulated accidents are dealt with more quickly, and that the plant is in a hazardous condition for less time.

REFERENCES

1. W. Petrick, K. Ng, C. Stuart, D. Cain. "A Production System for Computerized Emergency Procedures Tracking." EPRI Seminar: Expert Systems Applications in Power Plants, Boston, Massachusetts, May 1987.

2. "Emergency Operating Procedures Tracking System." EPRI Report NP-5250M, June 1987.

3. W. Petrick, K. Ng, C. Stuart, D. Shen, D. G. Cain, and B. K. H. Sun. "Emergency Procedures Tracking System for Nuclear Power Plants", ANS Topical Meeting on Artificial Intelligence and Other Innovative Computer Applications in the Nuclear Industry, Snowbird, Utah, August 1987.

4. A. J. Spurgin, D. D. Orvis, D. G. Cain, C. C. Yau. "Testing an Expert System: Testing the Emergency Operating Procedures Tracking System," IEEE Fourth Conference on Human Factors and Power Plants, Monterey, California, June 1988.

5. G. W. Hannaman, A. J. Spurgin, Y. D. Lukic and D. Worledge. "Impact of a Behavior Model of the Control-Room Crew on Plant Transient Analyses," International Conference on Power Plant Simulation, Cuernavaca, Mexico, 1985.

6. G. W. Hannaman, A. J. Spurgin, Y. D. Lukic. Human Cognitive Reliability Model for PRA Analysis. NUS-4531, Electric Power Research Institute, 1984.

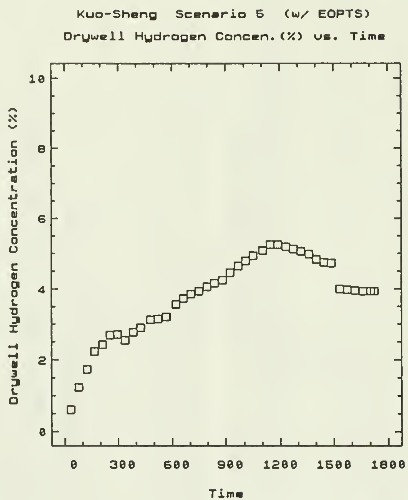


Figure 1a. Drywell Hydrogen Concentration, Scenario 5, EOPTS Crew A/B

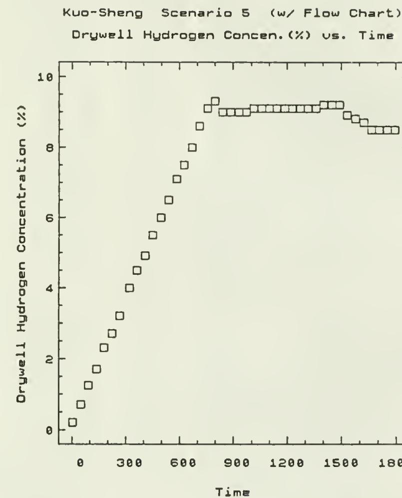


Figure 1b. Drywell Hydrogen Concentration Scenario 5, EOPS Crew E/A

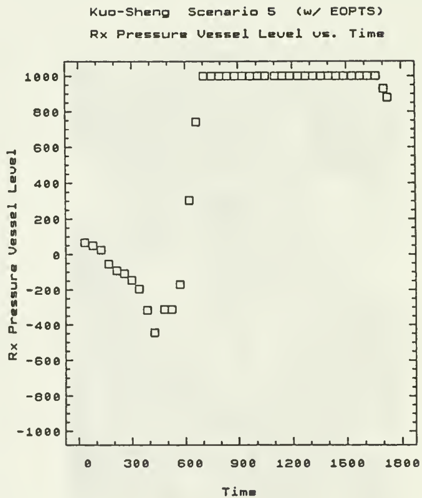


Figure 2a. RPV Level, Scenario 5,
EOPTS Crew A/B

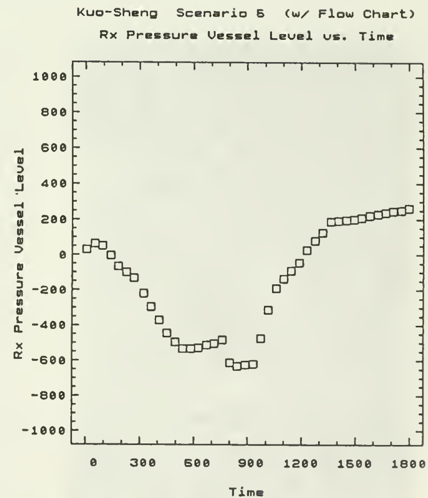


Figure 2b. RPV Level, Scenario 5,
EOPS Crew E/A

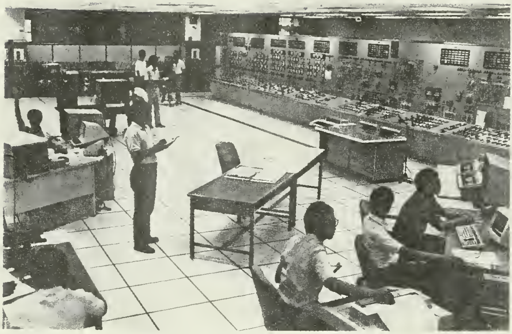


Figure 3. EOPTS Test Setup at Taipower's KuoSheng BWR Simulator Site; Observers in Place, Operation Crews Standby, EOPTS Display at Various CRTs, and Transient About to Start.

Figure 4. Crews using EOPTS Operate in the Rule-Based Cognitive Domain.

Figure 5. Crews Using EOP Flow Charts Operate in the Knowledge-Based Cognitive Mode.

Distributed Expert System Architecture Using a Dedicated Knowledge Server: An Innovative Solution for REALM On-Line

STEPHEN A. TROVATO and BOWIN M. LINDGREN
Consolidated Edison Co. of New York, Inc.
New York City, New York, USA

ROBERT A. TOUCHTON
Technology Applications, Inc.
6621 Southpoint Drive North, Suite 310
Jacksonville, Florida 32216, USA

ABSTRACT

This paper presents an up-to-date look at REALM, the Reactor Emergency Action Level Monitor Expert Advisor System, including recent innovations in the system architecture and our approach to Verification and Validation (V&V). The emergency classification domain is reviewed and the problem, solution and benefits are outlined. A REALM system description is then presented, followed by a description of the REALM V&V approach. The paper concludes with a look at how REALM is being generalized to embrace plant sensor interpretation beyond emergency classification (e.g. On-line Tech Spec or thermal performance monitoring) under the name of OASYS, for On-line Advisory SYStem.

EMERGENCY CLASSIFICATION DOMAIN BACKGROUND

For abnormal situations in a nuclear power plant where there is the potential for a significant release of radioactivity to the environment, the NRC requires that the utility owner of the plant have an emergency response plan to protect the health and safety of the public.

The NRC has established guidelines for utilities to follow which require that as part of the response plan, the utility develop a procedure to classify the level of severity of an event into what is called an Emergency Action Level (EAL). These emergency action levels are a kind of alarm to warn the NRC and state and local authorities of a serious problem.

There are four emergency action levels:

Notification of an Unusual Event - A variety of non-severe events that could signal the start of a potential problem. For example, something that exceeds the plant technical specifications (which defines the envelope for normal operations), or an earthquake or fire, or even the injury of a worker.

Alert - There is a degradation in the plant systems which could result in a significant release of radiation to the environment.

Site Area Emergency - Further degradation of plant systems to the point where a significant release is probable.

General Emergency - A significant release is occurring or has occurred.

In the unlikely event that an emergency situation were to arise at a nuclear plant, the operations staff would refer primarily to 2 sets of procedures:

Emergency Operating Procedures - which state how to restore the plant to a safe or normal condition.

Emergency Classification Procedure - which states how to assess the situation and classify the event into one of the four Emergency Action Levels.

These procedures are keyed to each other and trigger activities by off-site authorities at the alert level.

STATEMENT OF PROBLEM

During an actual event, the primary responsibility of the operations staff is to restore the plant to a safe condition in order to protect the public as well as plant equipment. The emergency classification process requires that the operations staff, particularly the shift technical advisor, turn his attention away from plant operation in order to interpret this procedure and perform the appropriate notifications of NRC and other authorities.

Determining the appropriate condition can be complicated because the determination about what conditions exist may require receiving and interpreting extensive information. For example, how does one know that the reactor coolant system is breached? There are many possible ways of this occurring. Also since there are many complicated rules that apply, interpretation can become difficult when a grey area is encountered. Interpretation may also vary depending on the shift crew.

Another aspect is the timeliness of notification. The NRC requires that the utility respond in a very short time, in some cases as quickly as 15 minutes. Under an actual event, operations personnel are swamped with alarms and information requiring their actions to control the plant. The event classification task is an extra burden which does not contribute to safe operation of the equipment.

A power company typically conducts an emergency drill for the NRC and several practice drills each year. In the past, some emergency classification calls have been made incorrectly or missed entirely during these drills.

THE SOLUTION FOR INDIAN POINT 2

At Consolidated Edison Company of New York, Inc.'s (Con Edison) Indian Point 2, the solution to the above problem is two fold. First, the site staff are making best efforts to simplify the procedures for emergency classification. This involves greater reliance on the state of the fission product barrier and less reliance on diagnosing specific events.

Second, the REALM expert system is being developed to provide the shift technical advisor with a tool that will provide advice well in advance of the time he will need it.

In 1985, the Electric Power Research Institute (EPRI) contracted Technology Applications, Inc. (TAI) to design and build an emergency classification expert system, now known as the REALM expert advisory system.

In 1986, Con Edison teamed up with EPRI and TAI as the host for developing an off-line prototype of the system. In 1988, the utility began the current research project to develop an <u>on-line</u> expert system, the first known attempt at such a system by a nuclear plant owner.

REALM is a good example of an "expert systems" application in that the emergency classification process requires inferencing on a great deal of information. The system is primarily intended as an aid to the shift technical advisor in the control room.

The success of REALM will be measured by its ability to provide a correct, consistent and most important timely response. The system can diagnose a condition significantly faster than a human. In use, it will already have reached a conclusion well before the shift technical advisor reaches the point in his procedures where he will need to consult it.

Another major objective is to provide a consistent method for emergency classification. The system will attempt to remove grey areas and provide a common mode of reasoning.

REALM BENEFITS

The system's primary benefit is its ability to provide expert advice when the expert is unavailable. REALM embodies the combined knowledge of a team of experts. This is another way in which an expert system can help. While the "experts" may be nearby, they may not be able to reach the scene in time or may not be able to give the task of emergency classification their full attention because their primary attention is the safe operation of the plant.

One side benefit is that improved diagnostic information on plant conditions will be made available to the shift technical advisor against which he can check some of the operations staff reasoning. It will enable him to check his thinking in a pressure situation (i.e., have I missed something?) and evaluate the consequences of his actions (i.e., if we take this component out of service will that put us into a higher emergency action level?).

The consequences of an incorrect classification are staggering. If the severity of an actual event is underestimated, the utility may not be taking the proper actions to resolve the problem and the utility could be fined by the NRC and be subject to the risk of law suits should public injury occur as a result. If overestimated, the more likely occurrence, it could cause an unnecessary mobilization of state and local emergency forces including, for example, moving 10,000 school children. Between the terrible publicity and the risk of injuries during such an event, public outcry would be devastating.

REALM will also be used as an aid during the 6 or 7 emergency drills held yearly. This use will provide a nearer term benefit, namely improving emergency drill performance, which will improve Con Edison's regulatory image, i.e. helping to achieve a better SALP (Systematic Assessment of Licensee Performance) rating.

REALM will document the decision making process and provide a trace or log of both events in the plant and reasoning by the operations staff. It will also be used to develop emergency scenarios upon which future drills will be based. Using

REALM to develop scenarios for future drills in house will save the company money and time.

Finally, it will be used to train personnel in emergency response. Using REALM for training will both improve the quality of training and again save money and time for training.

REALM SYSTEM DESCRIPTION

The primary function of REALM is to provide a prompt and accurate assessment of plant status with little or no operator input. REALM will provide expert advisories to Operations, Emergency Planning and Technical Support personnel in the identification and classification of emergencies and abnormal situations. The REALM expert system can be viewed as a collection of knowledge in the form of LISP program code, decision rules, and software objects grouped into knowledge bases.

Inputs and Outputs

At Indian Point 2, REALM will normally receive all the data it needs from the Safety Parameter Display System computer, which at Con Edison is known as SAS (Safety Assessment System). This system provides the operations staff with information on the critical safety functions which must be maintained. REALM relies primarily on the SAS computer for valid data. However, in many cases, REALM goes well beyond SAS both in attempts to test if valid sensor data is received and also to reach conclusions when data is invalid or missing. This is primarily achieved through its multiple reasoning paths.

A small amount of data for REALM will be manually input. This is primarily true when there is an observable condition; for example, "the containment hatch is open." REALM also allows the operator to override data known to be suspect if correct data is obtained from a locally read instrument.

REALM's principle output is a conclusion - the emergency action level. REALM reaches intermediate conclusions which identify plant conditions or states even though these may not be an emergency action level. For example, "Rapid Secondary Side Depressurization" has occurred. REALM provides a trace of the reasoning it used to reach its conclusion. REALM also allows the operator to propose questions like "What if?" For example, "What if another component cooling pump fails?" REALM gives the operator the ability to test the vulnerability to a given event. For example, Feeder 4A is the only one left that is supplying vital power. If it is lost, the condition will call for an escalation to "Alert."

REALM Functions and Features

REALM provides seven modes of operation at the RMTs: "On Line - Display", "On Line - Trial", "Off Line - Playback", "Off Line - Trial", "Off Line - Scenario Development", "Off Line - Training", and "Off Line - Curator" modes.
The first two modes ("On Line - Display", and "On Line - Trial") are on-line modes and will be used to monitor the actual plant by requesting the REALM computer's findings. The remaining modes are off-line and will be used for testing, support and model maintenance. When in one of the off-line modes, the system will read simulated data from the microcomputer's local data storage device (hard disk). The man-machine interface for all modes will be similar, with only a few differences reflecting the primary function of each mode. REALM provides the following modes and features:

On-Line Display mode - the user is made aware of the plant situation and emergency classification recommendation via visual and audible annunciations. In addition, the following features are provided:

Rationale Window - provides an English-language report explaining the system's current recommendation and underlying logic.

Response Display - provides a time-stamped English-language log of all interpretations, conclusions, and response to changes in plant conditions. A summary report lists the state of any off-normal conditions or threats.

Vulnerability Window - provides an English-language report of conditions or events which would cause the declaration of a more degraded situation.

Request Display - where REALM posts requests for situation-specific (i.e., sensor-driven) manual data. This would, in turn, free the user from having to decipher large amounts of manual data and focuses requested data to items that are pertinent to the current state of the plant.

Tabular Display - provides dynamic, on-screen tables indicating current state of data and knowledge. These tables can be printed or saved to disk.

On-Line Trial mode - the user has complete access to all sensor and manual data, thereby allowing the investigation of the consequences of changing plant operation ("what-iffing"). When this mode is entered, the Trial Mode inherits On-Line Display Mode data for that instant in time. Processing of On-Line Mode and Trial Mode continue completely in parallel until Trial Mode is exited.

Curator mode - It is expected that the REALM models will continue to evolve owing to changes in the plant design, procedures and industry regulations, and the discovery of additional knowledge that can be used to improve the plant model. As such, the custodian (the person authorized to modify REALM) of the system has been provided with an impressive collection of tools which make the maintenance and re-validation of the system as reliable and as efficient as possible. The Curator mode automatically generates hardcopy tables and diagrams which document the system's knowledge bases and rule bases, including interrelationships of objects and rules. Changes are recorded in a file so that an audit trail is available as a permanent record.

Playback mode - provides a testing and demonstration environment which fully emulates the On-Line Display mode using scenario files stored on disk.

Training mode - provides training in the interpretation of sensor data by playing back scenarios and allowing the trainee to compare answers with the "expert."

Scenario Development mode - facilitates the creation of test, demonstration, and training scenarios.

REALM Distributed Hardware and Software Architecture

The on-line REALM expert system will operate on a VAX and a network of COMPAQ 386 computers with a minimum of 12 Megabytes of Random Access Memory (RAM). The current REALM Architecture actually distributes the expert system processing demands by having a MicroVAX 3500 computer process and interpret the incoming data and a network of Compaq Deskpro 386/20 Remote Microcomputer Terminals (RMTs) display results and process operator requests for local analysis and evaluation of findings. Each RMT is, in fact, a full-scope REALM expert system, including the knowledge and rule bases.

Thus, the central REALM computer performs all primary REALM processing: data pre-processing, data evaluation by the REALM expert system, and communication of the findings to the RMTs. The RMTs each independently provide the user-demanded features of REALM: explanation facility, vulnerability analysis, trial mode, response log and tabular and printed reports. This means that each user can be exercising any of the available features without any impact on the performance of the other RMTs or the central REALM computer. RMTs are currently slated for the central control room, the technical support center, the emergency operations facility, the emergency planning offices and headquarters (Manhatten).

The portions of the system residing on the VAX are written in a combination of DEC's VAX Common LISP and VAX C. The operating system is VMS. The portions of the system residing on the COMPAQ 386 are written in a combination of Golden Common LISP (a version of the LISP language produced by Gold Hill Computers, Inc.) and Microsoft C. The RMTs use DECnet DOS to communicate with the MicroVAX 3500 computer over an Ethernet link. The REALM knowledge bases, rule bases and user interface are written in the KEYSTONE expert system development environment.

The REALM man-machine interface is resides on an RMT configured to require minimal operator training and operator interaction when operating in the on-line modes, including on-screen prompting and context-sensitive help screens. This is accomplished by incorporating state-of-the-art human factors capabilities such as color images, cursor pointing and selecting devices and pop-up menus. The interface uses a cursor pointing device (mouse or trackball) for rapid cursor positioning and item selection. The design of the man-machine interface was designed to conform to current human engineering guidelines such as Computer-Generated Display System Guidelines (EPRI NP3701). Three of the users will be able to control REALM (that is override data) while two of the users will have a read only link. Only one remote terminal will have control at a time under password control.

REALM Concept of Operation

Incoming data is collected and processed by the generic pre-processor module and placed in "objects" within the expert system knowledge bases. The central process will then cause the REALM experts to "inference" on the changed data. "Findings" will be placed back into the knowledge base "objects" and will be available to the other rule based experts (Figure 1). REALM then broadcasts its conclusions to the network in order to update the various RMTs.

REALM's assessment of the plant relies on a hybrid architecture and uses both rule-based reasoning and object-oriented programming techniques. The REALM environment represents (as "objects" within the knowledge bases) the Indian Point 2 power plant instruments, systems and sub-systems, components, accidents, events, conditions, statuses, and resources as required to support decision-making. The decision-making knowledge is represented in rule bases and consists of two general classes: "event-based" rules, which strive to determine the presence of predefined events, and "symptom-based" rules, which strive to provide meaningful findings even when no specific problem events can be identified. Rules may be explicitly based on source documentation, such as background documents and operating procedures, while other rules may be more heuristic in nature, relying on operator experience or engineering judgement for justification. The REALM concept is structured to model the reasoning process used by each domain expert and therefore incorporates a "team of rule based experts" approach. It is also designed to handle a well-behaved situation quickly and accurately using a minimum set of

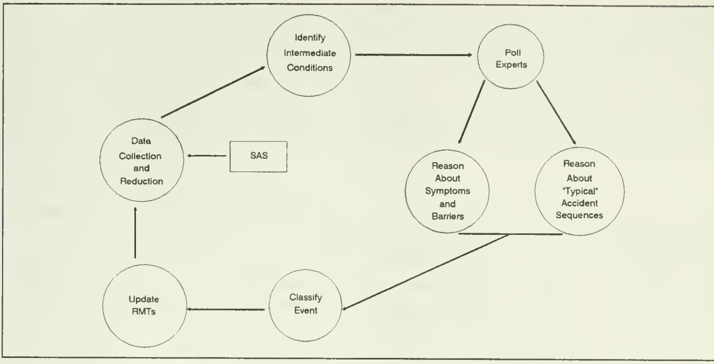


Figure 1. Reasoning Process

reasoning and resources. At the same time, it is prepared to handle a situation with missing or conflicting data and still arrive at the best possible conclusion using its team of rule based experts.

REALM VERIFICATION AND VALIDATION

Verification and validation of expert systems has been a concern, because multiple reasoning paths could create conflicts and are difficult to test in the manner that a conventional software system would be tested - input, process, output.

For REALM, we have taken a unique approach which we believe demonstrates that verification and validation of an expert system is actually easier than a conventional system.

The first step in developing an expert system is the knowledge engineering effort. During this step an attempt is made to capture expertise for a known domain. In our case, REALM, this involved review of the applicable plant documents (Emergency Operating Procedures, Emergency Classification Procedure, Technical Specifications, Final Safety Analysis Report, Abnormal Operating Instructions, Station Operation Procedures, Station Administrative Orders, NRC Guidelines and the Code of Federal Regulations) and interviewing plant staff (Operations, Safety Assessment, Regulatory Affairs, System Engineering and Emergency Planning). The key to the success of this step is to have a knowledge engineer (the person gathering the information) who is himself an expert in the domain.

The next step added specifically for this project was a decision model design review. We asked ourselves the question "What is different about this system that makes it so difficult to verify?" REALM reasons; it contains a complicated method of combining facts and rules in a manner that emulates the actual process performed by the shift technical advisor. But this actual process was defined by an engineer or team of engineers who understand the operation and response of the plant under abnormal conditions whether these are single or multiple failures. Therefore, the simple step needed to verify that REALM "thinks" correctly and provides correct advice is to review the logic of the system in the same fashion that engineers review the design of a plant system. Namely, add a series of design review meetings where the knowledge engineer presents his logic to a team of experts and together this group reaches agreement on the correctness of the system's reasoning. This is an application of standard engineering practices to a new situation.

One key to this step is that expert system shells provide features which make this process easy. Rules can be printed in a graphical diagram which shows how they link together; objects can be printed in a hierarchical diagrams which show their interrelationships, and the rules and descriptions can be written in a near English form which allows an expert with no computer background to understand how the information is represented in the software. Another key to this step is the design review process which brings together the combined knowledge of a team of experts to reach a consistent philosophy. This process actually resulted in improvements in the existing emergency classification procedures.

After this we apply standard tests to check the system.

Verification - Is the system being correctly designed to perform the intended task - Are we doing the right job?

Validation - Now that the system is built, is it working as we intended - Are we doing the job right?

OASYS = REALM - EALs

The software architecture developed for REALM was designed with a long-term general view of on-line expert advisory systems. Much of the underlying technology is common to all on-line situation assessment and analysis systems. Now that the Indian Point 2 REALM system is maturing, TAI is recasting the generic aspects of REALM as the On-Line Advisory System (OASYS). This expands the applicability of this powerful technology beyond that of emergency action level classification alone. In this light, REALM can then be considered as an application "instance" of OASYS.

The OASYS/REALM architecture is modular and expandable. The generic interface to on-line sensor data (e.g., SPDS) can provide an integrated environment (Figure 2) for EALs, Tech Specs, and thermal performance, or a variety of status monitoring settings. In whatever setting, the OASYS/REALM infrastructure (e.g., explanation facility, vulnerability analysis, trial mode, reports, tables, CURATOR mode, etc.) and methodologies (e.g., representation of instruments, diagnosis of system states, etc.) are substantially re-usable. Likewise, the development of OASYS/REALM to date has surmounted many technical problems associated with evaluating and analyzing live data on-line:

- temporal reasoning
- dynamic agenda
- generic interface/preprocessor
- distributed architecture.

Recall that REALM (and thus OASYS) is designed in a modular fashion and is based on an architecture comprised of a team of experts. The "team members" are in fact rule classes that reason upon plant components and instruments, as well as the findings of other "experts," modeled as objects in the knowledge bases. A new expert can easily be added to the system.

Con Edison, EPRI and TAI have expended considerable resources for the development and implementation of this system. Continuing to build on this technology will greatly decrease the technical risk to utilities embarking along these lines by leveraging off of this industry investment.

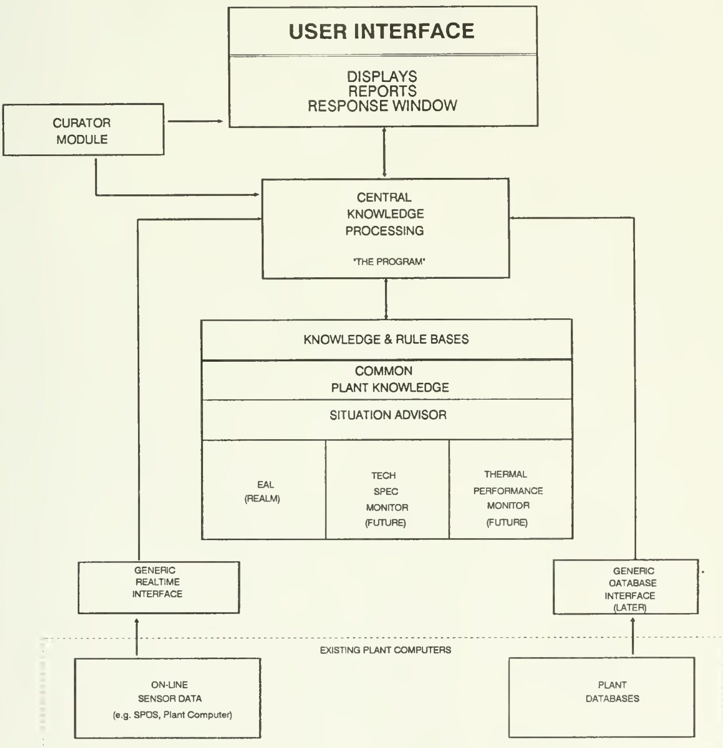


Figure 2. OASYS Architecture

TECHNOLOGY, TOOLS, AND METHODS

SELEXPERT: An Expert Advisor for Evaluating Candidate Expert System Projects

E. R. CREAMER, R. B. FRAHM, E. HYMAN, L. W. KAUFER, H. MAYER, H. T. ROMAN, and R. S. WITKOWSKI
Public Service Electric and Gas Company
P.O. Box 570 - 16G
Newark, New Jersey 07101, USA

ABSTRACT

Working with the results of several technology assessments performed by outside consultants, members of Public Service Electric and Gas (PSE&G) Company's interdepartmental artificial intelligence (AI) task force developed their own expert system for evaluating potential expert system applications. Named SELEXPERT by the group, the system was aimed at helping PSE&G employees to learn and understand basic concepts involved in expert systems design and application.

This paper will discuss PSE&G's experience with SELEXPERT, including specifically:

1) PSE&G AI Task Force activities as a prelude to development of SELEXPERT;

2) The SELEXPERT rule base and how it works;

3) Modeling considerations pertaining to the development of SELEXPERT.

PSE&G AI TASK FORCE ACTIVITIES

In order to understand the technical and economic implications of expert systems, and to determine where such systems might be used in the Company, PSE&G established an interdepartmental AI Task Force (1) in late 1985. The first meeting of the group took place in December 1985, with a Phase I report issued in August 1986. Phase I activities involved identifying potential applications, evaluating the state-of-the-art of AI technology, and determining the level of AI support in the public and private sectors. A Phase Ia report followed in December of 1986, which surveyed the AI vendor market for utility related expert system applications suitable for demonstration at PSE&G. Phase II activities involved screening potential applications for prototypical development. Phase II was completed in August 1988, and utilized two consultants, Texas Instruments and AGS, Inc. These consultants also provided valuable "knowledge engineer" training for selected task force members. Figure 1 illustrates the activities of the Task force.

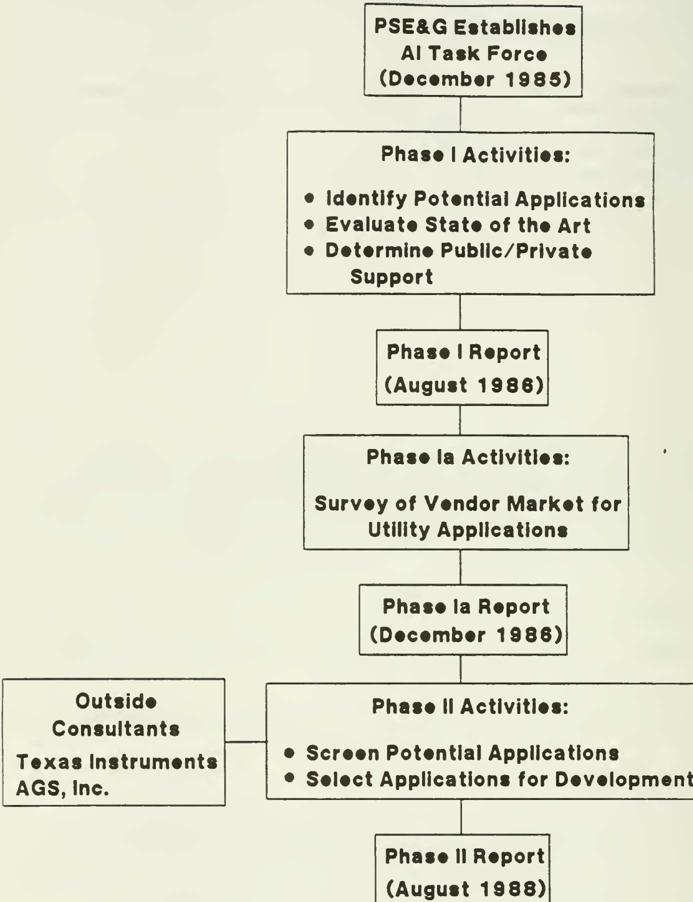


Figure 1

Activities of the PSE&G AI Task Force

Phase I identified 99 potential expert system applications which the task force grouped into similar families of applications. Of the identified applications, the task force selected 25 for detailed study and evaluation. With the assistance of Texas Instruments and AGS Inc., these applications were ranked and prioritized. Figure 2 lists the ranking of the selected applications by PSE&G department. The task force experience further contributed to a significant vision of the power plant of the future (2).

TASK FORCE SUBCOMMITTEE ON DOMAIN EVALUATION

Working in parallel with the consultants, PSE&G selected members of the task force and assembled them into a subcommittee charged with independently developing criteria for the evaluation of candidate expert system applications.

The intent in creating the subcommittee was to increase task force learning about the application evaluation process, as well as to provide an independent check on the consultant's work. Task force members' backgrounds included the engineering, research, library science, and information systems disciplines.

In preparation for their effort, several subcommittee members attended a three day course in Symbolic Processing presented by the consultant. The training proved invaluable in providing a technical foundation for later subcommittee tasks.

Drawing heavily on a commercially available training kit and an industry publication, the subcommittee developed a list of 24 True/False questions which could be used to evaluate a potential application. The questions were qualified as being related to either "business" or "technical" concerns including issues successful of value, appropriateness, and development.

Having completed development of their own set of evaluation criteria, and, impressed with a scoring scheme utilized in one of the consultants preliminary reports, the subcommittee decided to develop a similar method for translating answers to the 24 True/False questions into a simple score which reflected the overall suitability of an application for development using expert systems technology.

The subcommittee also decided to extend the scope of their effort to include development of materials which would assist potential PSE&G users of expert systems in:

1) Learning basic principles of expert systems and the expert system application evaluation process;

2) Proceeding with serious expert system development.

To extend the learning experience, the subcommittee decided that the knowledge acquired by the subcommittee should be incorporated into an expert system if possible. It was thought that development of such a system could also enhance transfer of the new technology to users.

Information Systems

Computer Equipment Operations	(1)
Network Troubleshooting	(2)
Help Desk	(3)

Human Resources:

Cut Score Evaluation	(1)
Management Job Evaluation	(2)
Grade B Job Evaluation	(3)
Career Path Recommendation	(4)

Nuclear:

Radiation Monitoring	(1)
Plant Chemistry	(2)
Electronic Diagnostics	(3) tie
Sequence of Events Analysis	(3) tie
Electronic Root Cause	(4)
Vibration Monitoring	(5)
Preventive Maintenance Scheduling	(6)
Mechanical Failure Analysis	(7)
Radiation Dose Analysis	(8)
10CFR 50.59 Evaluations	(9)

Fossil:

Power Brokering	(1)
Plant Chemistry	(2)
Sequence of Events (SOE) Alarm Analysis	(3)
Vibration Monitoring	(4)
Thermal Performance	(5)
Pump Failure Analysis	(6)
Computer System Troubleshooting	(7)
HVAC Problem Analysis	(8)

Note: Ranking is (1) being highest

Figure 2

PSE&G Department Ranking of Twenty-five Selected Applications

Adopting the "prototype" approach to system development frequently used in expert system development, one subcommittee member with some representational modeling experience took on the task of developing an automated scoring scheme. The system was tentatively named "SELEXPERT", meaning EXPERT system for the SELection of potential applications. A basic rule base shell, which had been purchased by the task force for earlier experimentation, was utilized in developing the prototype.

The system was patterned along the lines of the consultant's evaluation scheme which had impressed the subcommittee as providing a simple picture of the suitability of an application for development. The prototype, as developed, fit in well with the consultant's scheme. Initial valication runs comparing scores to those obtained by the consultant looked good. It was accordingly agreed to produce a basic expert system as a task force deliverable and to also translate the prototype into a manual scoring scheme which could be used by "computerphobes".

Following prototyping, a member of the subcommittee with experience in use of another, cheaper rule base shell suggested that SELEXPERT be rewritten using a shell which permits unlimited run time copies. The second shell also was viewed as being somewhat easier to use for beginners than the previous product.

SELEXPERT was shifted with little effort (much of the work was performed by a wordprocessing person given a "crash" course in the shell editor). The subcommittee also decided to make complimentary copies of the shell available to interested parties through the Research & development Department. A copy of the rule base runtime compiler was also purchased to allow delivery of a SELEXPERT version whose heuristics (and hence performance) could not be "damaged" by beginning users.

Later, during efforts to validate the use of the SELEXPERT, a Lotus 1-2-3 (TM) version was also developed and is now available to "spreadsheet" users. Seeing the potential utility in such an application, the PSE&G Information Systems Department has also decided to investigate development/acquisition of a more serious applications ranking product to be used professionally in departmental expert system development activities.

SELEXPERT - AN OVERVIEW

This next section of the paper focuses on SELEXPERT itself: what is does, how it was built, and how it actually operates. A number of actual screen displays are included to suggest the feel of the system and its operations.

As previously mentioned, SELEXPERT was designed to provide a basic score for a candidate application which would indicate the suitability of the application for development using an expert system. A broad group of users was targeted for the product, including:

1) An expert trying to gain insight into whether or not an expert system might be used to automate a task or problem in his/her area of expertise;

2) A manager trying to understand just what expert systems are all about. (a line supervisor at PSE&G was observed to remark following a expert systems indoctrination presentation: "Looks like something out of 2001 to me!").

3) Anyone with an interest in basic expert systems, how they work or how they are developed.

The present version of SELEXPERT was developed using Version 1.2 of the VP-Expert (TM) Rule Based Expert System Development Tool, from Paperback Software International. The final product was compiled for delivery at "runtime" using Version 2.02 of the VP-Expert (TM) Runtime Compiler. In addition to the features of the product as designed, any of the VP-Expert (TM) capabilities available in the runtime compilation may also be used (such as "why" or "what if" queries).

To avoid any complications due to misunderstandings about the degree of sophistication of the product or the purpose for its development, SELEXPERT was distributed for internal PSE&G use only and not for profit. The rule base documentation in SELEXPERT, as well as separate hard copy user documentation provided with the product, include disclaimers indicating the limitations of the product.

SELEXPERT was constructed to operate on either an IBM XT, AT or PS2 personal computer set up with the DOS and 640K of RAM; the system was made available on either 5.25" or 3.5" diskettes.

Reflecting the approach of the task force subcommittee, the representational model encoded into SELEXPERT was built to provide individual scores for each of eight criteria relating to the likelihood of successful development. Criteria scores are in turn rolled up into business and technical scores for the potential application.

Probably the best way to get a feel for how SELEXPERT works is to run through a typical consultation. The number of the figure illustrating the corresponding screen display is indicated in parentheses. Upon starting the consultation by entering the runtime command and the name of the application, the computer displays the SELEXPERT system header (see Figure 3).

A brief introduction is followed by simple instructions for using the system. The menu of applicable consultation commands is displayed below the consultation frame. It should be noted that more complete instructions for both SELEXPERT and VP-Expert (TM) features are provided in the accompanying hard copy user documentation.

Pressing any key prompts SELEXPERT to ask for the name of the application being evaluated and the date of the evaluation. These attributes are used if a hard copy printout is requested after the consultation. After the name and date are entered, SELEXPERT brings up the first of the 24 questions into the consultation frame (see Figure 4).

```
                         SELEXPERT
                        Version 1.0
                           1988

        Developed by the PSE&G Artificial Intelligence Task
          Force Ad-Hoc Subcommittee on Domain Evaluation

             Public Service Electric and Gas Company


Welcome to SELEXPERT, an expert system which provides advice
concerning the Selection and evaluation of potential EXPERT
system applications.

To evaluate a potential expert system application, indicate whether
the statements made by SELEXPERT about the application are
True or False (T or F).      (Press Any Key to Continue)
```

Figure 3

Initial SELEXPERT Display

```
To evaluate a potential expert system application, indicate whether
the statements made by SELEXPERT about the application are
True or False (T or F).      (Press Any Key to Continue)



Enter the name of the application being evaluated.
 Radiation Monitoring

Enter today's date.
 04-03-89


The application supports the CORE of the business?

(The task is essential to the creation of Corporate products and services, or
to the process of delivering them to the customer.)

 T                        F
```

Figure 4

First SELEXPERT Question To User

As is true of all questions, the possible choices in answering are displayed in a menu below the questions (in this case T or F for True or False). Additional information to assist the user in answering the question is provided in parentheses after the question, and the name of related variable is indicated by capitalization in the question text.

The user selects a response and enters RETURN. SELEXPERT stores the response then brings the next question into the consultation frame. Each additional question is in turn brought up after the user responds to the previous question, until all of the questions SELEXPERT needs to complete the consultation are unanswered. (Typical questions are illustrated in Figure 5.)

SELEXPERT only asks the questions necessary to evaluate the proposed application, parsing the rule base of any questions which are answered or preempted by responses to previous questions. The responses to previous questions, as well as any scores assigned to evaluation criteria, are withheld until the consultation is completed to avoid biasing the user.

Upon completing the consultation, SELEXPERT displays the results of evaluating the application, including individual criteria scores and final scores for both the business and technical aspects of development. Criteria are grouped with the aspect to which they apply (for example, the criterion Management is under the Business section). All scores are presented in terms of the intuitive and often used "1 to 10" scale.

Pressing any key (Figure 6) causes the system to inquire as to the user's preference for output, either None or the printouts displayed in Figures 7 and 8. Printouts of the evaluations scores, consultation answers, or both may be selected. Printouts include the name of the proposed application and the date of the consultation, useful for historical documentation purposes.

During a consultation, the various VP-Expert (TM) "Go commands" may be used to display additional information concerning a particular question or conclusion. For instance, selecting "How" will display information about "how a conclusion was reached". The user chooses the variable of interest from a list of the names of user choice, intermediate or conclusion variables, and the reason for the value of the variable is displayed. If the variable was set by the user, the system displays "because: You said so.".

Selecting "Why" on the other hand displays the reason the question currently under consideration in the consultation was asked. "How" and "Why" are related through VP-Expert through the "BECAUSE" statement of explanation which the programmer has attached to a given rule. For instance, the answer to a query "Why" a question is asked is the "because" attached to the rule which fired the question. The answer to "How" a factor variable was set is the "because" attached to the rule which set the variable or, if user set, "because: You said so.".

Another VP-Expert (TM) feature available during the consultation is "?" response for "unknown". This feature allows the user to respond that the value of a variable or a answer to a question is unknown. If the answer to

```
Development is within the current expert systems STATE-OF-THE-ART?
Has a system performing a similar type of task been developed elsewhere?

(Due to the nature of the knowledge processing involved, some tasks may be
more difficult to capture in an expert system than others, and previous
experience with a similar application may be helpful.  AI Task Force contacts
can help you with the types of tasks to which expert systems may be applied,
as well as a list of specific systems which have been developed.)

 T                        F


The task is can be classified as NARROW and self-contained?

(The aim is to select a limited task within the domain.  The task should
be defined very clearly and should be of a step-by-step nature.  The task
should not involve either diverse sources of knowledge or numerous
interdependencies with other activities/tasks.  This question is required to
take into account PSE&G's currently limited experience.)

 T                         F
```

Figure 5

Sample Technical Factors Questions To User

```
                         ************************
                            EVALUATION RESULTS
                         ************************

                      CRITERION          SCORE

                      Impact             = 7
                      Payback            = 9
                      Constraints        = 7
                      Management         = 7

                      Total Business Score = 7.545455

                      Expertise          = 9
                      User               = 7
                      Knowledge          = 9
                      Task               = 9

                      Total Technical Score = 8.750000

                                (Press Any Key to Continue)
```

Figure 6

Sample SELEXPERT Score Display

```
                Payback            = 9
                Constraints        = 7
                Management         = 7

                Total Business Score = 7.545455

                Expertise          = 9
                User               = 7
                Knowledge          = 9
                Task               = 9

                Total Technical Score = 8.750000

                         (Press Any Key to Continue)



Indicate the printout desired (if any):
 None                     Scores                    Answers
 Both
```

Figure 6a

```
*************************
   EVALUATION RESULTS
*************************
```

APPLICATION: Radiation Monitoring   DATE: 04-03-89

CRITERION	SCORE
Impact	= 7
Payback	= 9
Constraints	= 7
Management	= 7

Total Business Score = 7.545455

Expertise	= 9
User	= 7
Knowledge	= 9
Task	= 9

Total Technical Score = 8.750000

Figure 7

SELEXPERT Printout Of Scores

****

QUESTIONS AND ANSWERS

****

APPLICATION: Radiation Monitoring DATE: 04-03-89

The application supports the CORE of the business? T

The application supports a Corporate/STRATEGIC objective? T

The application supports a SCARCE expertise in the user environment? T

The application either displaces costs, adds VALUE, or supports a strategy in the process? T

The need for the task will CONTINUE for several years? T

An improved UNDERSTANDING of the problem gained through expert system development will be valuable to the organization? T

The potential impact of the IMPRECISION of expert systems on the business is understood? T

The use of an expert system will not be politically sensitive or CONTROVERSIAL? T

There is an influential CHAMPION? Strong managerial support? T

There is a strong SPONSOR organization? T

At least one practicing domain EXPERT can be identified? T

The expert can COMMIT sufficient time to the project? T

The expert is ENTHUSIASTIC about the project? T

The expert possesses good COMMUNICATION skills? T

The user understands LIMITATIONS of expert systems and can live with them? T

The user group is COOPERATIVE and patient, and they have agreed to support the project? T

Performing the task for which the expert system is being considered primarily requires SYMBOLIC reasoning rather than numeric computation? T

Figure 8

SELEXPERT Printout Of Question Responses any questions related to a criterion is "?", SELEXPERT scores the related criterion at 5. This allows continuing the consultation, with a median value being used to evaluate the application.

Final remarks about use of SELEXPERT include the fact that the Lotus 1-2-3 version may be used to get a better view of the system workings, with variable valued being visible throughout a consultation and changing as individual questions are answered. Alternatively, the VP-Expert shell may be used to enter the SELEXPERT rule base and directly edit the system, although changing the rules will affect the performance of the system in terms of validity.

After a consultation using the shell is completed, the user may query "What if" a variable value is changed. The system will provice the variable list, and will reevaluate the application using any new values provided for variables. If a "what if" variable is the answers to one of the 24 questions, the system will reask the question and any related questions triggered by the new response provided. Values for criteria scores may be reassigned directly when prompted by the system "What is the value of (variable)?".

SYSTEM DESIGN CONSIDERATIONS

Since SELEXPERT, as well as most expert systems, involves a significant amount of representation (heuristics represent knowledge), it seems appropriate to discuss some of the modeling considerations used in the design of the system.

It has also been the experience of some of the PSE&G AI task force members that the lack of understanding of representation and the related art of modeling have been an obstacle to understanding expert systems and their application. Related to the previous problem, the thinking that conventional systems may be equally well used for development of applications involving the processing of knowledge has been observed.

The effort to design SELEXPERT supported the idea that representational modeling concepts are important to expert system design. Unfortunately, these concepts are not centralized in any single discipline, with a number of different related paradigms in existence. The addition of the expert system, and more recently the expert support system (ESS) concepts further cloud the issue. In any case, continued development, documentation, and dissemination of the experience and theory of representation is needec.

Turning to the specifics of the SELEXPERT design effort, the general considerations involved included:

1) The basic model design;

2) The model structure;

3) The scoring scheme;

4) Model verification and validation.

Basic Model Design

Probably the most important decision involved in the design of the SELEXPERT was to produce a small, simple model based on "deep" knowledge of the evaluation process. The applicable principles here were to build a "robust" and "parsimonious" model capable of performing well in a very broad user domain and simple enough to enable a beginning user to gain understanding of expert systems and the domain evaluation process.

The nature of the task, which would be classified generally as involving "interpretation" of information/knowledge about the potential application, and to some extent "prediction" of the likelihood of success in undertaking development, was not optimally matched with a rule based approach. However, it was felt that by keeping the model simple and working within the flexibility of the rule based concept, a satisfactory representation could be constructed.

Fringe benefits of this approach were that using a rule base shell was within the limited skills of task force members, and building a simple system allowed keeping the total number of rules well under 100, thus eliminating any performance problems when delivered on widely available conventional P/Cs.

Overall, the model concept then was one of a "top-down" representation incorporating expert knowledge about domain evaluation. In addition to providing a "general" user interface due to the scope of potential users, the user was maintained in the system to provide needed expertise and knowledge concerning the various evaluation factors (hence the product should probably be rightly termed an ESS).

Attributes of the system that came with the development approach included the fact the system would be 100% correct due to the use of heuristics, and that the user would be likely to gain the benefits of increased learning and understanding that normally accrue with use of a representational model.

Model Structure

On of the more important principles used in the area of modeling is that, all other factors being equal, a model which parallels the structure of the reality being modeled would be expected to perform in a superior manner to one which did not. Although it is not clear that theory is well established here, one might explain this in terms of gaining overall "validity", and hence lower level "replicative" and "predictive" validity, by incorporating high level "structural" validity directly into the model.

The incorporation of structural validity also adds robustness and parsimony to the model, due to the stability and better fit provided by the high level theory involved. Parsimony probably most importantly supports increased robustness by eliminating unnecessary and burdensome aspects of the model. When the representation involves significant complexity, robustness in itself become an important design objective.

A third benefit of this approach in this case was that the understanding of expert systems and their evaluation would be enhanced by a structurally valid model, particularly if the user looked into the system as an example of an expert system itself.

Structural validity, robustness and parsimony may be obtained in a number of ways, most of them "tricks" of the modelling art. Probably the most straightforward way is to build proven relationships or methods directly into the model. Features of SELEXPERT design reflecting this principal include the use of existing commercial products and publications as the basis of the questionnaire, and patterning the evaluation process after that used by a successful knowledge engineering firm.

Other more detailed aspects of this approach utilized in the design of the system structure included the following:

1) The 24 questions were selected by the subcommittee to represent basic fundamentals of domain evaluation. The level of subcommittee understanding was probably suited to abstraction of these fundamentals (whereas experts may have made the model too complicated).

2) Evaluation criteria were developed based on intuitive constructs affecting development success and the various questions were then discretely related to the criteria.

   - This provided a structurally valid decomposition and needed decoupling.

   - The criteria fell generally in line with the consultants', supporting their validity and providing a convenient means of validating the underlying model.

3) Evaluation scores were combined into either a Business or Technology composite score using weighting factors and a weighted average.

   - This separation reflected the original thinking of the group, and allowed the user to focus on the less familiar technical concepts.

   - The weighting factors allowed adjustment of the model to changes in the business environment and provided some "modeler controlled" variables which could be used to fine tune the model without altering the basic structure.

   - The alignment with the consultant's model allowed using the consultant's weighting factors, reflecting their expertise and providing a starting point for tine tuning the model.

<u>Scoring Scheme</u>

The principles involved in the development of the scoring scheme parallel those for the evaluation criteria and incorporate several additional concepts. Basic thoughts employed in design of the scoring system included:

1) The True/False (essentially bipolar) format for the questions was used to force the user to make a decision concerning a factor, to add robustness given the range of users, and to provide needed variance reduction.

2) Unique criterion scores were assigned to different combinations of question answers as follows:

   - The 1 to 10 scale was adopted because it was simple, intuitive, and well known

   - Even scores (2, 4, 6, 8, 10) were deleted as a variance reduction measure

   - A score of 0 was assigned if an "essential" factor was not present reflecting the subcommittee thinking

   - Factor interrelationships were assessed to determine the proper score for a criterion (for example, whether they were conditional, independent, or mutual)

   - Values of 3, 5, and 7 were used for the general span of scores, 1 and 9 for extreme situations

   - The discrete combinations were adopted overall as structurally valid representations of factor/criterion relationships and to add variance reduction

Verification and Validation

Verification of SELEXPERT was performed informally through the review of the system by subcommittee members and other interested PSE&G individuals during development. Diskettes of the product were distributed allowing on-line verification. The parallel development and review of the manual scoring scheme was also useful in verifying the design.

Although technically a verification issue, the validation of the underlying model received more formal attention. Even though exceptional performance was not seen as essential, good performance gave needed reassurance that the subcommittees thinking was on track.

Reflecting the goals in building the model, validation focused on assessing whether or not the model "replicated" the evaluation process, and further generally predicted the suitability of an application for development.

Since the model strongly paralleled existing methodologies, verification provided adequate validation of replicative validity. Predictive validity was largely assessed by comparing scores with those independently obtained by the consultant. Additional applications whose general suitability to development were mutually agreed to be subcommittee members were also evaluated.

Figure 9 summarizes the validation runs and shows surprisingly good performance by the model. Incidentally, SELEXPERT itself evaluated well as an application (although interpretation of this fact is left to the reader!).

Finally, some efforts were made to validate SELEXPERT from the user perspective. These generally consisted of review of the product by subcommittee members, as a diverse group of semi-knowledgeable users; less knowledgeable but "friendly" users were also exposed to the product in several instances. Any comments from use of SELEXPERT were discussed by the subcommittee members and appropriate changes made to the system or documentation. Work on the text of the questions, and particularly the related additional information, is ongoing.

CONCLUSIONS

PSE&G's artificial intelligence task force captured its own knowledge, acquired from consultants and during its three years of work, in SELEXPERT, an expert advisor which evaluates proposed expert system applications. This working product successfully models a consultant's evaluation process. Both SELEXPERT itself and the story of its creation will be useful in training others to properly understand the design and use of expert systems. SELEXPERT has also pointed to the value of a more sophisticated tool for use by the Information Services group at PSE&G as a "knowledge-engineering advisor", and efforts are under way towards this end.

ACKNOWLEDGEMENTS

PSE&G wishes to acknowledge the work of its contractors, Texas Instruments Incorporated and AGS Incorporated (New York), and also two articles which contributed to the SELEXPERT concept and to the list of twenty-four questions, references 3 and 4 below.

REFERENCES

1. H. T. Roman, F. A. Marian, R. J. Mears, E. A. Hyman, "Expert Systems Applications at PSE&G - A Task Force Approach", presented at the EPRI Conference on Expert Systems in Power Plants, Boston, 1987.

2. H. T. Roman, F. A. Marian, "The Year 2000 Power Plant", IEEE Computer Applications In Power, April, 1989.

3. David Prerau, "Selection Of An Appropriate Domain For An Expert System", A. I. Magazine, Summer, 1985.

4. "The One-Minute Engineer", a questionnaire included in the text of a knowledge engineering course offered by Mind Path Technologies of Dallas, Texas.

POTENTIAL APPLICATION	CONSULT SCORE	SUBCOMM SCORE	SELEXPERT BUS	SELEXPERT TECH	SELEXPERT AVG
SEQUENCE OF EVENTS	5.7	- -	7.5	4.4	6.0
VIBRATION MONITORING	4.8	- -	7.5	3.1	5.3
PREV MAINTENANCE	4.7	- -	7.2	1.9	4.5
MECH FAILURE ANAL	4.6	- -	4.1	5.6	4.9
PLANT CHEMISTRY	6.4	- -	7.5	5.4	6.5
ELECTR DIAGNOSTICS	5.7	- -	5.2	5.6	5.4
ELECTR ROOT CAUSE	5.5	- -	7.5	4.4	6.0
RAD DOSE ANAL	4.4	- -	5.5	4.4	5.0
RAD MONITORING	7.5	- -	7.5	8.8	8.2
10CFR50.59	4.0	- -	4.5	4.1	4.3
SELEXPERT	- -	7.0	7.2	7.9	7.5

Figure 9

SELEXPERT Model Validation Results

A Verification and Validation Methodology for Expert Systems in Nuclear Power Applications

DANIEL B. KIRK
The MITRE Corporation
Civil Systems Division
1820 Dolley Madison Blvd
McLean, Virginia 22102, USA

JOSEPH A. NASER
Nuclear Power Division
Electric Power Research Institute
3412 Hillview Avenue
Palo Alto, California 94304, USA

ABSTRACT

The potential for expert system applications in the nuclear power industry is widely recognized. The benefits of these systems include the retention of specialized human expertise, improved equipment reliability through enhanced diagnostics, and consistency of reasoning during off-normal situations when operators are under great stress. However, before any of these benefits can be realized in critical nuclear power applications a careful and comprehensive Verification and Validation (V&V) program must be applied to ensure the quality of the application.

This paper provides a summary of a methodology for the V&V of expert systems developed for nuclear power applications. The similarities and differences of expert system and conventional software techniques are identified and analyzed, and conventional V&V approaches are advocated where applicable. When the conventional approach is not applicable, V&V techniques specific to expert systems are presented and integrated with conventional methodologies to form a disciplined methodology suitable for nuclear power applications. This methodology is tailored to each of various types of expert systems, where the types are defined according to the difficulty of performing V&V on each type. These guidelines must be further tailored to the unique features and uses of each expert system developed for a particular nuclear power application.

1.0 INTRODUCTION

Verification and Validation (V&V) is an essential activity for software which performs critical activities such as those found in nuclear power plant applications. Due to its importance in ensuring the quality of the product, V&V has been used extensively in the Nuclear Power Industry to ensure software quality. Examples include on-line systems such as the Safety Parameter Display System (SPDS; Straker, 1981) and analysis tools such as the RETRAN thermal-hydraulic code (McFadden, et al., 1987).

Expert systems have a great potential for application in the Nuclear Power Industry; however, they cannot be exempted from the requirement for a complete and through V&V program, particularly if they are to shift from their current use in a primarily advisory mode to that of a controlling function. The benefits of expert systems include consistency of reasoning during off-normal situations when humans are under great stress, the reduction of time required to perform certain functions, the detection of incipient equipment failures through predictive diagnostics, and the retention of human expertise in performing specialized functions. As these potential benefits are demonstrated and realized, the development of expert systems will become a necessary part of the Nuclear Power Industry. To this end, the Electric Power Research Institute (EPRI) has launched a broad-based exploration of potential expert system applications intended to augment the diagnostic and decision-making capabilities of personnel. The goals of this effort are to enhance safety, human productivity, reliability, and performance (Naser, 1988). Two examples of existing systems are the Emergency Operating Procedures (EOP) Tracking System (Petrick and Ng, 1987) and the Reactor Emergency Action Level Monitor (REALM) System (Touchton, 1988).

An obstacle to the acceptance of expert systems is the lack of a methodology for their V&V. The V&V of expert systems is not a straightforward task. They differ from conventional software in several respects, and so a conventional software V&V methodology cannot be directly applied to their V&V. For example, expert systems employ rules with a declarative, rather than procedural, representation and so do not always follow simple procedural steps. Also, expert systems often follow a cyclic development process rather than the straight-line path of conventional systems. These differences cause problems that require special attention. There are, however, also many similarities and analogies with conventional software and its design process that can help in devising methods suitable for expert systems.

This paper provides a summary of a methodology for the V&V of expert systems developed for nuclear power applications [a more complete description of this approach may be found in two EPRI reports "Approaches to the Verification and Validation of Expert Systems for Nuclear Power Plants" (Groundwater et al., 1987) and "Verification and Validation of Expert Systems for Nuclear Power Applications" (Kirk and Murray, 1988); the current paper draws heavily on this latter publication]. In this methodology, the similarities and differences of expert system and conventional software techniques are identified and analyzed, and conventional V&V approaches are advocated where applicable. When the conventional approach cannot be applied, V&V techniques specific to expert systems are presented and integrated with conventional methodologies to suggest a methodology suitable for nuclear power applications. This methodology is tailored to each of various types of expert systems, where the types are defined according to the difficulty of performing V&V on each type. These guidelines must be further tailored to the unique features and uses of each expert system developed for a particular nuclear power application.

Conventional software V&V was chosen as starting point for this expert system V&V methodology because the benefits of the conventional approach (for example, the emphasis on a requirements document) has been demonstrated numerous times in a wide variety of systems. The generic usefulness of such features, coupled with the criticality of nuclear power applications, argues that the burden of proof regarding the inclusion/exclusion of conventional components in a expert system V&V methodology be with those advocating their omission.

Before proceeding with a description of the expert system V&V methodology, it is useful to first define two terms. The first of these is that of "V&V" itself, so that there will be a clear definition as to the meaning and purpose of V&V. The second such term is that of "expert systems"; the definition used here is broader (and the resulting V&V methodology more comprehensive) than that used by some authors. A good deal of the vagueness and disarray associated with current views on expert system V&V can be traced to the variety of definitions available or to the flexibility of interpretation of these definitions.

2.0 DEFINITIONS

2.1 V&V

Following (Deutsch, 1982), verification may be defined as an activity that ensures that the results of successive steps in the software development cycle correctly embrace the intentions of the previous step. Each level of specification and the deliverable code are traced to a superior specification; i.e., the specification or code is verified to ensure that it fully and exclusively implements the requirements of its superior specification.

Also following (Deutsch, 1982), software validation may be defined as an activity that ensures that the software end item product contains the features and performance attributes prescribed by its requirements specification. It is important to note there that the software end item product does not necessarily refer to the final, deliverable code: in the structured design process which a good V&V program will enforce, the software will be designed in modules. Each of these modules should be individually validated against their own set of requirements as should, of course, the complete software program. Also note that testing of both the complete program and its modules is included in the validation effort. Testing is part the process of ensuring that the software end item product contains the features and performance attributes prescribed by its requirements specification.

Typically the above-defined term "software validation" will be simply referred to as "validation." There is a second kind of validation that is of importance here, namely that of requirements validation. This form of validation - also a portion of V&V activities - is the process of ensuring that the process of translating the customer's operational needs into an explicit set of software requirements has been done correctly.

2.2 Expert System

The term "expert system" has a variety of definitions. We shall adopt one here that covers a broad range of systems that others might call "knowledgeable" but not "expert" (cf. Waterman, 1986). We define an expert system to be any computer program for solving problems by using a rule-based approach. The system may contain procedural code or other forms of knowledge organized in tables, databases, etc., but it always must be based at least partly on a knowledge base that consists of a set of rules and facts. For that reason, "knowledge-based system" is an alternative, and sometimes preferred, name. Another alternative is that of "production system."

3.0 CONVENTIONAL V&V SOFTWARE METHODOLOGY OVERVIEW

The V&V of conventional software programs is a well-established and mature discipline. A description of this methodology is given in (Groundwater et al., 1987) and (Kirk and Murray, 1988); a more detailed treatment may be found in (DeMarco,1979) and (Deutsch, 1982). These references also describe the linear, stepwise, system lifecycle - otherwise known as top down design, or the waterfall method - that is used in the conventional V&V approach. This lifecycle, along with associated V&V activities, is illustrated in Figure 1.

Corresponding to the above V&V definition, V&V activities may be broken into three categories: 1) Requirements Validation, 2) Verification, and 3) Validation of the software system. Prior to the initiation of any formal V&V activities, a V&V plan should be submitted to the customer for approval. This plan, the Software Verification and Validation Plan (SVVP) should describe the methods (e.g., inspection, analysis, demonstration, or test) to be used to:

1) Validate the Software Requirements Specification (SRS),

2) Verify that:

   (a) The requirements in the SRS are implemented in the design expressed in the Software Design Document (SDD),
   (b) The design expressed in the SDD is implemented in the code, and

3) Validate that the code, when executed, complies with the requirements expressed in the SRS.

This plan is critical in that it forces the V&V team to plan their efforts and is the primary means of communicating these plans to the customer for review. The plan will typically be modified throughout the course of the software project as modifications and further specifications of future V&V activities are made. ANSI/IEE Standard 1012-1986 provides excellent guidelines for the construction of the SVVP.

Following the approval of the V&V Plan, requirements validation is the first formal V&V activity. This effort is probably the most critical V&V effort as the validated requirements document (the SRS) will form the basis for nearly all further V&V activities. Requirements validation is typically accomplished by a constructive approach such as data flow diagrams (DeMarco, 1978). This approach is constructive in that it provides both a method for constructing the requirements and a graphical method for clearly displaying the requirements to aid in their validation. The goal of requirements validation is to ensure that the requirements specifications (the SRS) is unambiguous, complete, verifiable, consistent, modifiable, and usable in operations and maintenance. The SRS must clearly and precisely describe each of the essential requirements (functions, performances, design constraints, and attributes) of the software and the external interfaces. Each requirement must be defined such that its achievement is capable of being objectively verified and validated by a prescribed method (eg., inspection analysis, demonstration, or test). A full discussion of the characteristics of a good requirements specification may be found in ANSI/IEEE 830-1984.

The second V&V activity is that of verification, which comes into play as more detailed system requirements are generated, and in the design process, as the System Design Description (the SDD) is produced. At each stage, the SDD must be verified to ensure

DEVELOPMENT STAGES	OUTPUT DOCUMENTS	BACKGROUND V&V ACTIVITY
CONCEPT PHASE	• Initial Software V&V Plan (SVVP)	• Establish V&V Plan
REQUIREMENTS PHASE	• Requirements Specification (SRS) • Validation Test Plans (System and Acceptance)	• Validate requirements [Evaluate requirements vs application Analyze interface requirements] • Generate System and Acceptance Test plans
DESIGN PHASE	• Design Description (SDD) • Validation Test Plans (Component and Integration)	• Trace SDD to SRS • Evaluate SDD • Generate Component and Integration Test Plans
IMPLEMENTATION PHASE		• Verify translation of design to code • Evaluate code • Generate test procedures
TEST PHASE		• Execute tests
INSTALLATION AND CHECKOUT PHASE	• V&V Final Report	• Audit installation • Prepare final report
OPERATION AND MAINTENANCE PHASE	• Updated SVVP	• Configuration Management

Figure 1. Conventional, Top-Down, Approach to Software System Development

that the document fully and exclusively implements the requirements of its superior specification (a full discussion of the characteristics of a good SDD is given in IEEE Standard 1016-1987). This activity of verifying the SDD is primarily a paper activity, i.e., that of comparing two sets of documents, but an important verification function is also aimed at facilitating the generation of these documents. To do this, the V&V team ensures that various requirements and design reviews - e.g., the Software Requirements Review (SRR) and the Preliminary Design Review (PDR) - are held to facilitate a review of the requirements/design specification and to encourage interaction between the various design team members. Further review and interaction is facilitated by assuring that design walkthroughs are held. These walkthroughs are informal meetings in which the author of a design product explains the details of the design to other members of the design team, the V&V team, and possibly the customer.

The final V&V activity is that of software validation. This goal of this effort is to validate that the code, when executed, complies with the requirements expressed in the SRS. As noted above, individual software modules - as well as the final, integrated software product and system - should be tested. This activity should begin in parallel with the requirements validation effort, so that as the system requirements become defined, explicit methods for testing those requirements are generated. This early emphasis in generating tests will help ensure that the requirements are indeed verifiable. Generation of tests should also occur throughout the verification efforts, so that as the system becomes more completely specified, more specific tests are generated. Tests should determine at a minimum: (a) compliance with all functional requirements as a complete software end item in the system environment, (b) performance at all hardware, software, user, and operator interfaces, (c) adequacy of user documentation, and (d) performance at boundary conditions and under stress conditions. ANSI/IEEE Standard 829-1983 gives excellent guidelines for the construction of a software test plan and test procedures. ANSI/IEEE Standard 1008-1987 gives similar guidelines for the testing of individual software modules.

4.0 DIFFERENCES OF EXPERT SYSTEM AND CONVENTIONAL SOFTWARE TECHNIQUES

The differences between expert system and conventional software techniques may be classified into two areas: 1) the differences between the software itself, and 2) the process by which the software is constructed (eg., differences in the software lifecycle phases).

4.1 Differences in Expert System and Conventional Software

Expert systems and conventional software differ in a variety of areas. The first difference between the two arises directly from the definition of an expert system; expert systems are constructed (at least in part) of a knowledge base consisting of rules and facts. This rule-based format allows an explicit representation of knowledge that has several benefits in V&V. The explicit representation makes that knowledge easier to understand and compare to the system requirements. In addition, it allows for various test for internal consistency and completeness of the knowledge base (Nguyen et al., 1987; Bonasso and Henke, 1988), and it often allows the use of an expert system building tool to apply that knowledge.

A second difference between expert systems and conventional software stems directly from the first difference - the declarative, rather than procedural, representation makes it difficult to implement conventional, structured design techniques such as those for tracing data flow (DeMarco, 1979). Such techniques rely on the decomposition of functional units into subunits, which in turn may be subdivided. This decomposition allows for the tracing of requirements to various levels of the system. Rules, however, have no structure for incorporating such a hierarchy, with the result that rules dealing with a number of different cases are often grouped together.

A third difference is that with expert systems there is often no single, correct answer for a given scenario. There may be a variety of acceptable answers as in, for example, configuration programs that shuffle fuel assemblies and inserts (Naser et al., 1987). If multiple correct answers are possible, then the V&V program must give special attention to criteria for determining correctness and comparison of alternative solutions.

A fourth difference that is related to the existence of multiple correct answer is the use of uncertainty in expert systems. The use of uncertainty can greatly complicate the V&V of expert systems because the number of possible logic paths greatly increases. In addition, the mechanism used for expressing uncertainty must be examined to determine that it allows an adequate representation of the actual uncertainty and properly propagates this uncertainty in the inferencing process.

The fifth difference between expert systems and conventional software is that the process which the conventional software performs - particularly for critical systems - is already often codified, i.e., there is a fixed set of procedures for carrying out the task that have already been approved. As will be discussed below, expert systems may also classified as "codified" in that they are based on codified knowledge, but typically expert systems - even for critical applications - are not based on codified knowledge. This knowledge must be obtained from experts through knowledge engineering and must be codified as part of the V&V process.

4.2 Differences in the Expert System and Conventional Software Construction Process

There are three principal differences in the expert system and conventional software construction processes. The first difference is that the knowledge base requirements and specifications for an expert system cannot, in many cases, be determined before knowledge engineering has begun in the design phase. Therefore, the complete validation of those requirements and specifications and the development of knowledge base test cases must be deferred to the design phase.

The second difference in the two construction processes is the rapid prototyping approach typically used in expert system construction. The rapid prototyping approach has both an advantage and a disadvantage with respect to V&V. The advantage is that the early prototypes provided by the rapid prototyping approach allow abbreviated V&V cycles to be completed early in the design phase. In particular, some validation of the prototype can be carried out to obtain to good estimate of the effectiveness/feasibility of the final system. In a conventional software approach, validation can only be performed after design and coding are complete. Software/performance defects found at this late stage are usually difficult to remedy. The disadvantage of the rapid prototyping approach is that the prototype is often transformed into the final system without the requisite V&V being performed. By the very nature of the rapid prototyping process, the prototype cannot be carefully V&V'd as it evolves. Simplifying assumptions, coding errors, poor documentation and a poorly structured system are often characteristics of a rapidly constructed prototype, and these are often best treated by simply discarding the prototype (which has served its purpose) and completely redesigning and recoding the system according to the conventional software construction process.

The final difference in the expert system and conventional software construction process is the use of an expert system building tool in the former process. This difference yields two points that relate to V&V. First, the expert system building tool can, and must, be V&V'd by conventional methods. If the tool has already been V&V'd, then this process need not be repeated for each individual application. The second point is that the building tool may suffice for prototype development, but it cannot 'scale up' to operation in deployment because of limitations not apparent to either the design or V&V team during the prototyping effort. The building tool must be evaluated very carefully before the prototyping effort begins (and constantly re-evaluated as that effort proceeds) for its suitability in the operational environment.

Using the above differences between expert systems and conventional software (and their development methodologies), it is possible to construct an expert system V&V methodology that is based upon conventional software V&V and addresses the special concerns of expert systems. Before outlining that methodology, it is first useful to classify expert systems into a number of types so that the V&V methodology may be tailored to those individual types.

5.0 EXPERT SYSTEM TYPES

The fact that expert systems vary in the source and type of knowledge stored or in whether uncertainty is explicitly recognized or not furnishes a convenient basis for classifying them. For example, the simplest expert system measured by these characteristics would be one that embodies straightforward coding of validated and verified decision tables. Its search space could be small, like all the possible choices in tic-tac-toe, and could be examined with exhaustive search techniques. Or, it could be large but factorable so that defined areas for the search space could be treated separately, and perhaps in an optimum sequence. Strategic guidelines would be (at least theoretically) available for narrowing the search and making it efficient. Even if every segment of the search space must be searched, the fact that it can be broken into pieces reduces each part to manageable size. Solving a succession of such minor problems can greatly decrease the total search time. Expert systems with such small or large but factorable search spaces will be termed "Simple." Those systems which are not simple are termed "Complex." These latter systems are primarily research systems. Included in this category are systems that employ such research issues as non-monotonic reasoning, multiple knowledge bases with potentially conflicting heuristics, or learning systems. Since these types of systems are still in the research phase, it is virtually impossible to make generalizations about their V&V at this time.

The dichotomization of expert systems into Simple and Complex categories may be further refined by splitting each of these categories into two sub-categories depending on whether or not the system incorporates in its design some method for handling uncertainty, i.e., uncertain information or uncertain logic. Uncertainty may apply to the existence or value of input conditions, the relationship of knowledge items or the validity of the rules. Such uncertainty can be made to reflect the expert's uncertainty of the input data, or the applicability of the rule to these antecedent conditions, or the appropriateness or certainty of the conclusions. Expert systems may embrace any of these forms of uncertainty, sometimes combining multiple uncertainties in reaching a result.

The characterization of expert systems may be still further refined with one additional discrimination - whether the expert system relies on previously codified knowledge or, conversely, relies on elicited (not previously-validated) knowledge. As discussed above, the validity of this latter (elicited) knowledge must be determined as part of the V&V process. Systems relying on previously validated knowledge are typically based on codified decision tables and thus fall into the Simple category of expert systems. As a result, this final factor only refines the Simple category of expert systems. The resulting 6 types of expert systems are shown in Table 1.

An example of a Type 1 expert system is the Emergency Operating Procedures (EOP) Tracking System (Petrick and Ng, 1987). The objective of this system was to develop an automated EOP tracking system that can first analyze nuclear plant conditions in real time and then identify appropriate emergency procedures and explain the rationale for taking them. It consists of a custom inference engine written in the "C" language for fast execution and a knowledge base of if-then procedures derived from the EOP guidelines developed by the BWR Owners Group. It is a Type 1 system because it relies on previously codified knowledge and does not use uncertainty. The V&V of this system is discussed in (Kirk and Murray,1988).

An example of a Type 3 expert system is the Reactor Emergency Action Level Monitor (REALM) System (Touchton, 1988). REALM is designed to provide real-time expert assistance in the identification of a nuclear power plant emergency situation and the determination of its severity. It has been structured to model an emergency classification process which might be used by the emergency director and his technical support group during an actual emergency. REALM consists of a number of distinct but interactive elements: interface, objects, "a team of experts," a series of message boards, and rules. The existence of multiple experts in REALM would seem to argue that it is a Complex type of expert system and thus very difficult to V&V. Fortunately, the multiple experts in REALM are partitioned into nearly disjoint functions, and thus may be considered a Simple type of expert system. Since REALM is based partly on elicited information and does not employ uncertainty values, it is a Type 3 system.

6.0 A V&V METHODOLOGY FOR EXPERT SYSTEMS

6.1 Establishing the System Requirements

The requirements document is a logical starting place for an expert system V&V methodology that is built upon conventional software V&V, as it is the central reference to all conventional software V&V activities. A requirements document should be written - or rewritten - whenever it is possible to do so, even though development, coding, or even testing, may be well under way. A clear statement and detailing of a system's requirements either demands or implies certain internal qualities of the software that can be affirmed by analysis and it provides external performance goals that can be explicitly affirmed by tests.

In some cases the requirements are known from the codified knowledge source or after sufficient effort is spent on eliciting expert knowledge. In other cases, where the development is gradual, consisting of alternating periods of incremental building and testing, requirements gradually emerge in better and more complete form as performance is making a similarly gradual improvement. The building of expert systems must often follow this cyclic, incremental, development pattern. The pattern

Table 1

EXPERT SYSTEM TYPES

TYPE NUMBER	DESCRIPTION
1	Simple, based on codified knowledge
2	Simple, as (1), but with uncertainty handling
3	Simple, based on elicited knowledge
4	Simple, as (3), but with uncertainty handling
5	Complex (generally for research)
6	Complex, as (5), but with uncertainty handling

Table 2

EXPERT SYSTEM CHARACTERISTICS, DESIGN GOALS, TEST CATEGORIES, AND/OR CANDIDATE REQUIREMENTS

CATEGORY	REQUIREMENT 1	REQUIREMENT 2 ...
Decision Quality, Correct Response		
Correct Reasons		
Usability 1. Ease of Use a. Interface b. Expertise Needed 2. Response Time		
Modifiability, Adaptability Reliability		

corresponds well to a model of development attributed to Boehm (Boehm, 1988) and is illustrated in Figure 2.

The cyclic model illustrates the position of requirements in the development cycle. At least a rudimentary notion of the requirements starts the first cycle. It steers the acquisition of knowledge and is gradually improved and enlarged as knowledge is acquired. Requirements development, as an accompaniment of knowledge acquisition, eventually enables expert knowledge about the application domain to be translated into facts, rules, or other knowledge representation structure. The process of translation starts with specifying the rules, etc., the hierarchy or structure, if any, within which they reside, and ends with the coding of a prototype system. Testing the prototype reveals deficiencies in performance, suggests holes in the knowledge base and stimulates another round of knowledge-building, coding, and testing.

In this cyclic model, requirements definition has a recurring role. This role can be implemented by pausing to formalize the requirements before each new round of coding begins. In general, for this or any other development cycles or patterns, the guidelines should be:

1. Strive for a requirements specification. If there is none, write one as soon as possible; improve it as further knowledge is gained about the application.

2. Let requirements specification interact with and be a partner of knowledge acquisition, as well as a guide to design. For these reasons, do not relegate requirements specification to an independent group, shutting out the designers.

3. Use requirements specification to guide the planning of validation tests and the identification of test criteria. Do this as early as possible, even though full-system testing must wait for the completion of coding and assembly. If a V&V team is to be used, get them started on test planning during the requirements analysis. Include designers on the V&V team.

4. Begin the planning of validation tests as early as requirements are available. Periodically consider whether and how requirements may be traced in the development stages. Can they be used as verification criteria in the translation from requirements to design specification, or from specification to coding?

There are several benefits to be gained from starting very early to try to formalize the requirements and from making an early start in planning validation tests based on those requirements. Awareness of the need for a requirements specification can help steer knowledge acquisition, and vice versa, as well as steer system design. Early planning of validation, based on requirements, sharpens the definition of what is wanted from the system and may stimulate the selection of verification tests to be applied as the system is being built. The careful examination of requirements, which is necessary for planning validation tests, may also benefit collecting and organizing the requirements themselves. In addition to these potential interactions, early validation activity promotes the early discovery of errors and omissions and the accompanying reduction in cost of remedying these errors.

6.1.1 Planning for System Validation. As just noted, an important component in establishing the system requirements is planning for the system validation. As the rapid prototyping approach will allow validation efforts to be applied early in the development cycle, the planning for the final system validation can also be a cyclic, evolving process. Apart from this difference in developing the system validation procedures and the specific concerns with validating the knowledge base (as discussed below) there will be little difference between the validation of an expert system as opposed to a conventional system. The primary questions that should be kept in mind as the validation process is being constructed are: What exactly should be tested? For whom is it being done? Who does it have to satisfy? What are the standards by which evaluations will be judged or scored? Above all, the overall guideline that must be followed is "write testable requirements/test to requirements."

As an aid to assuring that important considerations are not left out of the specification process or the evaluation process, it is desirable to generate a list of candidate qualities or capabilities to be considered. Even before anything much is known about the detailed aims of the project, it is likely that a candidate list of requirements subjects can be composed. To keep track of such subjects and help insure that they are addressed in formal requirements, a table of design goals, much like Table 2, can be helpful, at least as a starting point. As information is obtained in knowledge elicitation, in prototype tests and elsewhere in the usual iterations of development, the requirements in each category can be filled in, or the categories can be modified if needed. The completed table can be filled in, or the categories can be modified if needed. This table can be viewed as either a guide to, or a summary of, the requirements specification.

6.1.1.1 Object-Oriented Programming as an Aid to Validation. An expert system's rule base is characterized by its declarative, rather than procedural, nature. Conventional (e.g., structured) design techniques, such as tracing the data flow in data flow diagrams, cannot be applied directly to this declarative form of the rule base. The use of object-oriented programming can alleviate that handicap and improve the reliability, maintainability and understanding of expert systems. The changes that object-oriented programming permit in expert system design can improve validation by making the program easier to compare to the system requirements.

Object-oriented programming (Pascoe, 1986) is a general concept that brings to expert system design essentially the same benefits that it provides to any software design. This programming technique organizes a program in terms of modules, where each module may be thought of as an object with its own set of applicable operations. Each object has its own means of communicating and interacting with other objects in the program, and each stores and manipulates data in its own private section of memory. An object response is triggered by a message passed to that object asking it to perform the operation on itself. The details of how it performs the operation, however, are private, and need not be known or addressed by the message. This characteristic of hiding details can make programming easier to do and to understand. Messages can be expressed in general terms such as "reduce flow by 10%;" any module receiving that message "knows" what detailed operations have to be performed to accomplish it and can go about doing it in its own particular, internally programmed, way. Object-oriented programming can also permit objects to inherit the attributes of other objects (eg., the process by which an object reduces flow), thus reducing the reducing the amount of code that needs to be programmed, validated and maintained.

Object-oriented programming may be combined with a rule-based approach; such an approach is exemplified in the Alarm Filtering System (Corsberg, 1986). In this system, objects are used to represent the alarms and alarm states. Rules represent the expert system's control and decision-making process. Because of the modularity and the ability to conceal within each module details of how the object behaves or operates, the rules can be generic and thus can address many types of objects. As a result, in this particular system there are only 30 rules. The simplicity conferred by the abstraction and inheritance properties of this type of programming allowed the number of alarms and states in the system to be increased from 80 to over 200 in less than two days.

6.1.1.2 Planning for Validation of the Knowledge Base. As with any software module, the knowledge base must be separately validated against its own set of requirements. Part of the requirements must, of course, be an objective test-based requirement in which assertions and conclusions are compared with those of an expert (preferably in a double-blind experimental setting). This type of requirement, while useful, is not specific to expert systems in that one is simply testing the output of the software module. The explicit, declarative nature of the knowledge base allows a rather different type of validation test in which one can "lift the hood" and have the expert and other members of the validation team inspect the internals of the knowledge base for correctness. There are several techniques that can be used to aid this process. As with other aspects of validation planning, these techniques should be considered early in the requirements specification process.

The first two of these techniques are aimed at making the knowledge base more understandable and accessible so that it can more easily be inspected for correctness and completeness. In the first of these techniques, rules are subdivided into rule-groups; the function of each of these rule groups is explicitly defined, as is the external interface of each rule group. This external interface will typically consist of the list of facts which, if asserted, can satisfy an antecedent of a rule in the group, and a list of facts which can be asserted by a rule in the group. Sets of rule-groups may be packaged together into a higher-level unit called a rule object. The rule object may be treated as any other object in an object-oriented system, with its own private section of memory and communication with other objects (which may also be rule objects) via messages. As with other objects, the rule objects are invoked by sending messages to and from other objects. Such a packaging allows a means of incorporating rule-based processing in an object-oriented system while still retaining all of the advantages of the object-oriented paradigm (cf. Section 6.1.1.1). The previously discussed Alarm Filtering System (Corsberg, 1986) is an example of a nuclear power-related system using the rule-object approach.

The second technique aimed at making the knowledge-base more understandable and accessible is to display the relationship between the predicates and objects in various rules in a graphical format (Bonasso and Henke, 1988). To enhance the understanding of the interdependence of the rule-base, the graph can be inspected by panning, highlighting, or selecting various subgraphs (eg., displaying only those predicates and objects associated with a given rule group). The method usually used here is to place each predicate involved in a rule at a node in the graph. A directed arrow between nodes indicates that one predicate is used to compute the value of another predicate. For example, if we have a rule to deduce in a backward-chaining manner that a cylinder is stuck as

```
if      air-supply( ?line_x, ?cylinder_x)
             and  hot( ?line_x)
            then   stuck-cylinder( ?cylinder_x)
```

and a forward chaining rule to determine if the variable ?line_x is an air-supply to the variable ?cylinder_x of

```
            if       carries-air( ?line_x)
              and  joins( ?line_x, ?cylinder_x)
              and   input( ?line_x, ?cylinder_x)
            then   air-supply( ?line_x, ?cylinder_x)
```

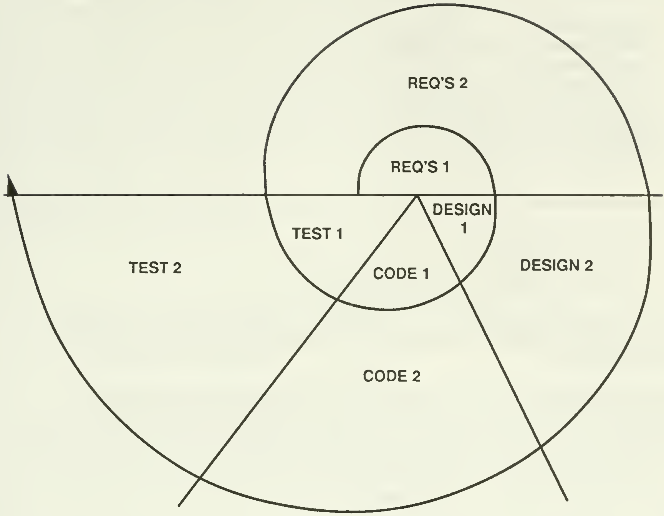
we can show the relationship of the predicates air-supply, hot, stuck-cylinder, carries-air, joins and input as shown in Figure 3. A similar graph for objects may be drawn for objects and object-classes referenced in rules.

The third technique involves generating a record of all the deductions that can be made for a given scenario input (Bonasso and Henke, 1988). This record can be inspected for correctness and completeness and can be used to help validate changes to the knowledge base. If such a record is made before and after modifications to the knowledge base, the difference between these two records can computed to allow a rapid identification of the differences induced due to the knowledge base modification.

6.2 Verification Issues Specific to Expert Systems

There are two verification issues that are specific to expert system V&V. The first of these is to ensure that the System Design Document completely and explicitly the describes the processing the expert system is to perform. The second of these is verifying the internal consistency and completeness of the knowledge base. The term "internal" is used here because we are not concerned with validating the correctness of the knowledge base against some external standard (e.g., comparing it against the expert's knowledge), but rather with the syntactical correctness of the knowledge base. Automated methods for checking the knowledge base internal consistency and completeness are somewhat analogous to the error-checking performed at compilation and run-time of conventional software.

6.2.1 The System Design Document. The System Design Document (SDD) for an expert system must address a number of design issues that are specific to these type of systems. First, all information that is input to the expert system must be described. This information must include the input source, the process or rule in the expert system requiring the information, and any restriction on the allowable range of the input. The SDD must also specify the set of facts that can derived during the inferencing process. If such an enumeration of these facts is not feasible, then the set of predicates associated with these facts must be specified, along with a description of the possible domain of objects for each predicate. For example (following the air-supply and cylinder example given in Section 6.1.1.2), it must be specified that air-supply, for a specified set of cylinders, is a predicate for which facts may be asserted during the inferencing process. The inferencing process(es) to be used must be explicitly defined, as must any escapes from those process(es). The mechanism for providing reasoning explanations (e.g., responses from "how" and "why" queries) must also be described. Finally, the mechanism for uncertainty handling, if any, must be described.



INCREMENTAL SYSTEM DEVELOPMENT

Figure 2. Iterative Model of Expert System Development (after Boehm)

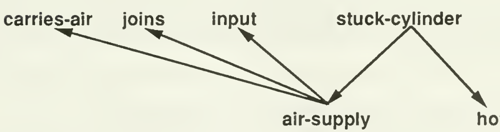


Figure 3. Predicate Graph Illustration

6.2.2 Verifying the Internal Consistency and Completeness of the Knowledge Base.
There are a variety of checks that can be performed to detect errors in the consistency and completeness of a knowledge base. These checks include consistency tests for

- redundant rules
- conflicting and potentially conflicting rules
- subsumed rules
- circular rules
- unnecessary if conditions
- illegal attribute values
- consistency of predicates
- consistency of variables

and completeness test for

- unreferenced attributes
- unreachable conclusions
- dead-end if conditions and dead-end goals

These tests are well-described in the current literature (Nguyen et al., 1985, 1987; Bonasso and Henke, 1988; Kirk and Murray, 1988; Stachowitz et al., 1988) and are not discussed further.

The above-listed consistency checks only detect problems in the knowledge base within individual rules and between pairs of rules, they cannot identify deeper inconsistencies that can arise during the inferencing process. Consider the following example taken from (Bonasso and Henke, 1988):

Suppose we have the following rules and facts:

```
if    (p) or (q)    then   (a)
if    (q) or (r)    then   (b)
if    (a)           then   (c)
if    (b)           then   (not (c))
(q)
```

There is an inconsistency in this knowledge base that would not be detected by any of the above-listed inconsistency tests: since (q) is true, then both (a) and (b) are true and so both (c) and (not(c)) are true, which is an inconsistent condition. Systems described by (Stachowitz et al., 1988) and (Bonasso and Henke, 1988) can detect these "deep" inconsistencies. However, due to the undecidability of first order (predicate calculus) logic, there can be no process to test for these inconsistencies that is guaranteed to terminate when an inconsistency does not exist. (Bonasso and Henke, 1988) have demonstrated that the removal of recursive rules and a restriction on the form of the knowledge employed can greatly reduce the chance of a non-termination, and have examined a method (termed lock resolution) which detects deep inconsistencies very efficiently.

7.0 INTEGRATING V&V INTO EXPERT SYSTEM DEVELOPMENT

As discussed in Section 5, expert systems vary in their complexity and their use of uncertain information and logic or their reliance on elicited knowledge. The kind of knowledge they contain and how that is obtained can affect not only the steps they go through in development, but also the kinds of errors that may occur. Systems that embody codified knowledge, such as decision tables extracted from an authoritative source, do not need iterative cycles of incremental development and can be designed very much like standard software, in a straightforward sequence of steps. Figure 4 shows a development scheme designed to fit this type of system. It allows for some recycling to reconsider the design if system tests reveal some deficiencies. Coding or design revision my also result from lessons learned in later, on-the-job, use of the system. Systems that implement knowledge elicited from domain experts often need the cyclical, iterative approach. Figure 5 shows a developmental life cycle that suits this type, allowing for linear development where possible but providing cyclical stages where necessary. Notes on Figures 4 and 5 indicate what V&V processes are relevant at the various stages of development.

A V&V program that fits the recursive style of expert system development may be summarized by the following activities:

- State the concept and tentative requirements.
- Collect expert knowledge and implicit requirements.
- Design and test the prototype system using the collected and engineered knowledge.
- Go back to collect more knowledge (and more rules and more identifiable requirements).

  The above steps may be repetitive, resulting in gradual enlargement and refinement of prototype(s) and performance. It usually results in gradual enlargement of the knowledge base.

- Review requirements list for accuracy, adequacy, completeness and attainability.
- Verify that requirements specification faithfully captures requirements, as listed.
- Verify - to the extent feasible - that the prototype design implements the requirements specification.
- Review the design for maintainability and modifiability. Consider the use of accounting such as dependency charts, or dictionary or directory tools (cf. Kirk and Murray, 1988, Section 6.3). Consider the maintainability/modifiability of the proposed architecture.
- Verify the adequacy and accuracy of how knowledge is represented in sensing, input, input processing and in the rules or reference data.
- Verify that all requirements are met at interfaces for which the project is responsible.
- Verify the internal consistency and completeness of the knowledge base.

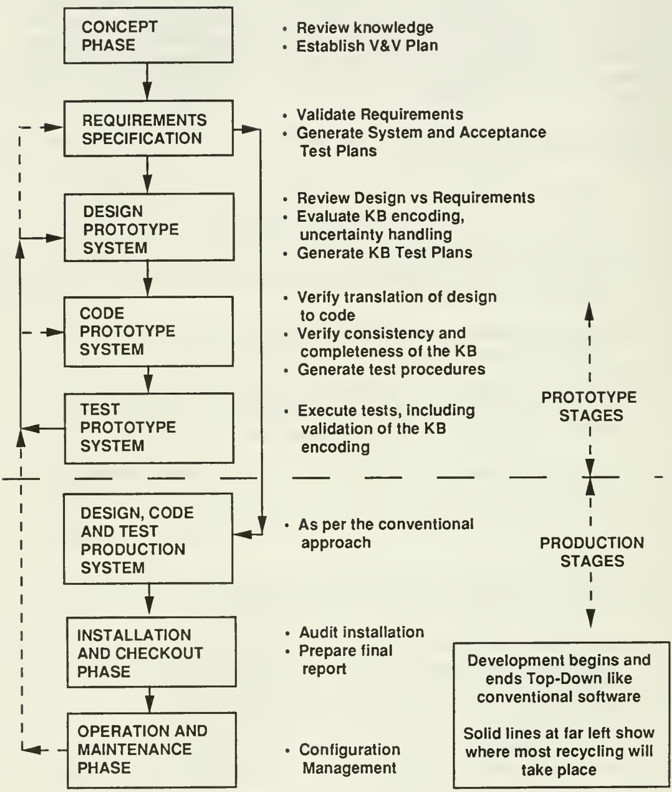


Figure 4. Lifecycle V&V of Expert Systems Embodying Only Validated, Codified, Knowledge

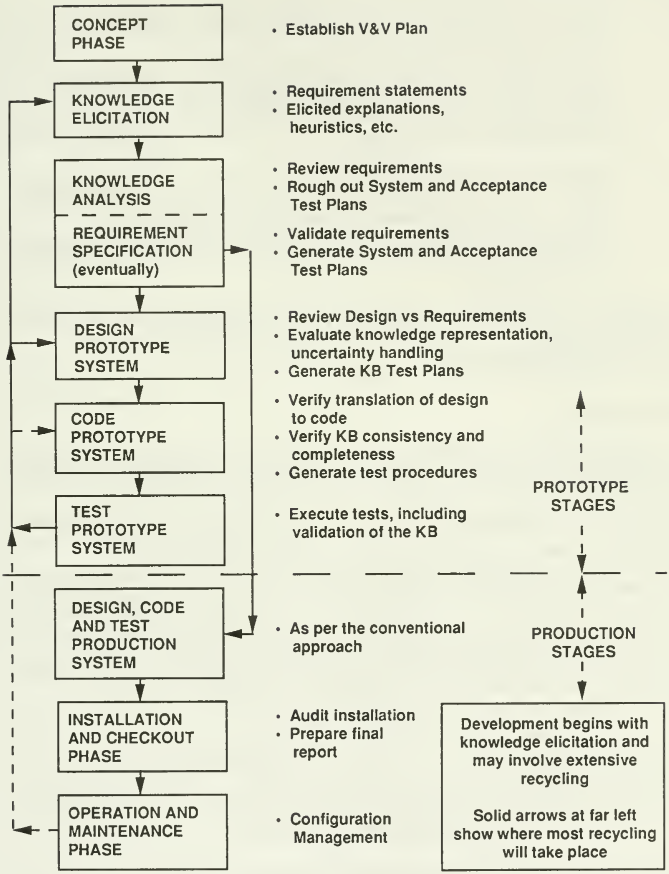


Figure 5. Lifecycle V&V of Expert Systems Embodying Elicited Knowledge

- Examine the knowledge base for correctness and the completeness of coverage of the domain. Consider the use of the knowledge base validation techniques discussed in Section 6.1.1.2.

- Conduct comprehensive system shakedown tests, exercising all inputs, outputs, decision path, etc.

- Verify usability, especially (but not exclusively) at the user interface. Employ subjective as well as objective criteria. The best policy is to include usability criteria in the system requirements and get users involved early for that purpose.

- Conduct selective tests, using carefully selected or designed special cases. Test on selected situations, scenarios, aimed to stress, explore, and bracket behavior. Test boundary conditions and thresholds. When incorrect behavior is detected, backtrack through the reasons and other antecedents of incorrect behavior, looking for the error source.

It is understood that any of the above steps may cause corrections to be made in some preceding design step(s). This recycling process is demonstrated by the feedback loops indicated in Figures 4 and 5.

8.0 CONCLUSIONS

V&V is an essential component of any system designed for critical applications such as those found in the Nuclear Power Industry. Expert systems have a great potential for application in this industry, but the lack of a methodology for their V&V is an obstacle to their deployment. This paper provides a summary of EPRI-sponsored work (Groundwater et al., 1987; Kirk and Murray, 1988) aimed at developing such a methodology. Although expert systems and conventional systems differ, it is suggested here that conventional V&V techniques be used as starting point for an expert system V&V methodology because of the solid track record and proven worth of the conventional techniques. With this starting point, the similarities and differences of expert system and conventional software techniques were identified and analyzed, and conventional V&V approaches were advocated where applicable. When the conventional approach was not applicable, V&V techniques specific to expert systems were presented and integrated with conventional methodologies to suggest a methodology suitable for nuclear power applications.

Expert systems were classified into six types to identify different V&V needs. Suggested methodologies were given for the first four types. The last two types of expert systems are still in the research phase and therefore it is not possible to identify appropriate V&V methods for these types at this time. V&V life-cycle activities for the first four expert system types are shown in Figures 4 and 5.

Additional work is being initiated to develop methodologies for nuclear plant V&V applications for knowledge certification and for developing validation scenarios. This work is being co-sponsored by EPRI and the Nuclear Regulatory Commission (NRC). The methodologies developed under this project will be tested on actual expert systems.

REFERENCES

1. ANSI/IEEE Standard 829-1983. "Standard for Software Test Documentation," New York, NY: IEEE.

2. ANSI/IEEE Standard 830-1984.,"IEEE Guide to Software Requirements Specification," New York, NY: IEEE.

3. ANSI/IEEE Standard 1008-1987, "IEEE Standard for Software Unit Testing," New York, NY: IEEE.

4. ANSI/IEEE Standard 1012-1986, "Standard for Software Verification and Validation Plans," New York, NY: IEEE.

5. ANSI/IEEE Standard 1016-1987, "Recommended Practice for Software Design Description," New York, NY: IEEE.

6. Boehm, B.W. (1988) "A Spiral Model of Software Development and Enhancement," Computer, Vol. 21(4), pp. 61-72.

7. Bonasso, R.P. and Henke, A.L. "KMAN: A System for the Specification and Management of Knowledge Bases," MTR-88W00187, McLean, VA: The MITRE Corporation.

8. Corsberg, D. (1986) "Extending An Object-Oriented Alarm-Filtering System," IEEE Expert Systems in Government Symposium, K. Karna, K. Parsaye, and B. Silverman, eds., McLean, VA. pp. 80-87: IEEE.

9. DeMarco, T. (1979) Structured Analysis and System Specification, Englewood Cliffs, NJ: Prentice-Hall.

10. Deutsch, M. (1982) Software Verification and Validation: Realistic Project Approaches, Engelwood Cliffs, NJ: Prentice-Hall

11. Groundwater, E.H. et al. (1987) "Approaches to the Verification and Validation of Expert Systems of Expert Systems for Nuclear Power Plants," NP-5236, Palo Alto, CA: The Electric Power Research Institute.

12. Kirk, D.B. and Murray, A.E. (1988) "Verification and Validation of Expert Systems for Nuclear Power Applications," NP-5978, Palo Alto, CA: The Electric Power Research Institute.

13. McFadden, J.H. et al. (1987) "RETRAN-02 - A Program for Transient Thermal-Hydraulic Analysis of Complex Fluid Flow Systems," NP-1850-CCM-A, Palo Alto, CA: The Electric Power Research Institute.

14. Naser, J. et al. (1987) "A Fuel Insert Shuffle Planner Expert System," EPRI Conference on Expert Systems in Power Plants, Boston, Mass: The Electric Power Research Institute.

15. Naser, J. (1988) "Nuclear Power Plant Expert System Verification and Validation," Validation and Testing Knowledge-based Systems Workshop, AAAI-88, Minneapolis, Minn: AAAI.

16. Nguyen, T. A. et al. (1985) "Checking an Expert System's Knowledge Base for Consistency and Completeness, " Proceedings of the Ninth International Joint Conference on Artificial Intelligence, Menlo Park, CA: AAAI.

17. Nguyen, T.A. et al. (1987) "Knowledge Base Verification," AI Magazine, Vol. 3(4), pp. 69-75.

18. Pascoe, G.A. (1986) "Elements of Object-Oriented Programming," Byte, August 1986.

19. Petrick, W. and Ng, K.B. (1987) "Emergency Operating Procedures Tracking System," NP-5250SP, Palo Alto, CA: The Electric Power Research Institute.

20. Stachowitz, R.A. et al (1988) "Research on Validation of Knowledge-Based Systems, "Validation and Testing Knowledge-based Systems Workshop, AAAI-88, Minneapolis, Minn: AAAI.

21. Straker, E.A., (1981) "Verification and Validation for Safety Parameter Display Systems," NSAC-29, Palo Alto, CA: The Nuclear Safety Analysis Center.

22. Touchton, R.A. (1988) "Reactor Emergency Action Level Monitor Expert System Prototype," NP-5719, Palo Alto, CA: The Electric Power Research Institute.

23. Waterman, D.A. (1986) A Guide to Expert Systems, Reading, MA: Addison-Wesley.

Human Factors Issues Related to Expert Systems for Electric Power Plants

RICHARD J. CARTER
Engineering Physics and Mathematics Division
Oak Ridge National Laboratory
P.O. Box 2008
Oak Ridge, Tennessee 37831, USA

ROBERT E. UHRIG
Instrumentation and Controls Division
Oak Ridge National Laboratory
P.O. Box 2008
Oak Ridge, Tennessee 37831, USA

ABSTRACT

The primary purpose of expert systems is to represent the knowledge of experts and make the expertise available to the human so that it can contribute to improved performance. In order to achieve this objective, human factors principles must be incorporated into the design. Two surveys oriented towards identifying the human factors issues related to expert systems were conducted. This paper describes the results from those surveys. It discusses the human factors issues under four main categories, the knowledge base of the expert system, the human-expert system interface, organizational support, and related topics (e.g., training, workload, and performance under stress). The viewpoints and opinions expressed herein are those of the authors and do not necessarily reflect the criteria, guidelines, and requirements of the United States (U.S.) Nuclear Regulatory Commission (NRC).

BACKGROUND

In the operation of an electric power plant, great quantities of numeric, symbolic, and quantitative information must be handled by the control room operator(s) even during routine operation. The sheer magnitude of the number of process parameters and systems interactions poses difficulties for the human, particularly during abnormal or emergency situations. Recovery from an upset situation depends upon the facility with which available raw data can be converted into and assimilated as meaningful information by the operator. Also, as in any complex sophisticated system operation, humans are sometimes affected by fatigue, stress, and environmental factors which in turn have varying degrees of influence on operator performance.

Expert systems are expected to take some of the uncertainty and guesswork out of the operator's decisions and to reduce his/her workload by providing expert advice and rapid access to a large information base. Application of expert systems to the control room activities in an electric power plant has the potential to reduce human error and improve plant safety and reliability. Furthermore, in a large number of nonoperating activities (e.g., testing, routine maintenance, outage planning, equipment diagnostics, fuel management, etc.) expert systems can increase the efficiency and effectiveness of overall plant and corporate operations.

Electric power utilities, equipment vendors, national laboratories, and consultants are developing expert systems for use in power plants. A number of these were presented at this and the earlier Electric Power Research Institute (EPRI) conferences on expert systems applications in power plants (1). The primary purpose of these expert systems is to acquire and represent the knowledge of experts and make the expertise available to the human so that it can contribute to improved performance. Hence, during the development of an expert system the interface between the human and the expert system should be optimized. In order to achieve this, human factors principles must be incorporated into the design. Unfortunately, until recently, the human factors issues related to expert system design, development, and implementation had not been fully identified.

RESEARCH PROGRAM

Oak Ridge National Laboratory (ORNL) is performing a research project for the U.S. NRC's Office of Nuclear Regulatory Research (RES). The overall objective of the project is to provide the technical basis for the development of regulatory criteria to evaluate the safety implications of human factors associated with digital and expert systems in nuclear power plants. One of the project's completed tasks was directed at the preparation of a program plan for regulatory, expert systems research. Another task was oriented towards determining the human factors issues related to the current, planned, and potential future uses of advanced instrumentation and controls, including expert systems, in the control room and technical support center.

As part of the development of the expert systems program plan discussions were held with sixteen NRC headquarters staff members, five from the RES, seven from the Office of Nuclear Reactor Regulation, three from the Office for Analysis and Evaluation of Operational Data, and one from the Executive Director's Office.

During the identification of the human factors/advanced instrumentation and controls issues, a survey of U.S. and Canadian vendors and utilities (i.e., United States - five utilities and five vendors, Canada - one utility and one vendor) was conducted.

The data collection instrument used during the NRC discussions was comprised of approximately twenty-five open-ended questions; the instrument for the utility/ vendor survey consisted of over eighty open-ended questions. The interviews were conducted by a team of two scientists, a human factors psychologist and a nuclear engineer with expertise in instrumentation, controls, and expert systems. Discussions at the NRC took place over a two-day period. The U.S. nuclear facilities were visited for one day each; the Canadian for a day-and-a half. Personnel at the NRC and each utility/vendor were interviewed either individually or in groups of two-to-five. The amount of time spent with particular people varied between one-half and three hours. Before each group of individuals was interviewed, they were informed of the purpose and background of the discussions/ survey and the benefits through their participation. They were told that their comments would be kept confidential and that no published material would identify remarks made by an individual or a specific utility/vendor. The data collection instruments were used to guide the course of the discussions and survey, but the interviews themselves were semi-structured and took form as they proceeded.

HUMAN FACTORS ISSUES

Human factors-expert systems issues, addressed in the program plan for regulatory research and identified during the survey of current, planned, and potential future uses of advanced instrumentation and controls, are exhibited in Table 1. A more elaborate presentation and discussion of the issues are described below. The human factors-expert systems issues have been organized under four main categories: knowledge base, human-expert system interface, organizational support, and related topics.

Knowledge Base

The knowledge base of the expert system contains the expertise (facts and heuristics), obtained either directly from experts or indirectly from books, publications, codes, standards, or data bases, as well as the general and specialized knowledge pertaining to the specific situation. The most powerful expert systems are those containing the most knowledge (2).

The correctness and completeness of the information within the knowledge base are the keys to obtaining reliable and valid solutions using expert systems. It is important to ensure that the knowledge base is also accurate and consistent. Two questions which must be addressed from a human-factors standpoint are: what are

Table 1. Human Factors Issues

Topics	Issues
Knowledge Base	o Adequacy of the Knowledge Base o Qualifications and Experience of the Expert(s) o Acquisition/Extraction of the Expert Knowledge o Knowledge Representation o Software Verification and Validation
Human-Expert System Interface	o Simplicity, Clarity, and Understandability o Support Effective Use o User's Perspectives and Mental Models o Explanation Facilities o User Friendliness o Mode of Interaction
Organizational Support	o Management Style and Support o Needs Assessment o Function Allocation and Division of Labor o User Involvement During the Life Cycle o Manner of Implementation o Use of Guidelines
Related Items	o Training o Impact on Workload o Effects of Stress o Performance Evaluation o Effect on Human Performance o User's Reaction o Over-Dependence

the tasks that the expert system is designed to perform and are they adequately represented in the knowledge base of the expert system?

A number of problems can exist in the knowledge base (3). They include: (a) excess generality or specificity [special cases overlooked or generality undetected], (b) concept poverty [useful relationship not detected and exploited], (c) invalid or ambiguous knowledge [misstatement of facts or approximations, or implicit dependencies not adequately articulated], (d) invalid reasoning [programmer incorrectly transforms knowledge], (e) inadequate integration [dependencies among multiple pieces of advice incompletely integrated], (f) limited horizon [consequences of recent, past, or probable future events not exploited], and (g) egocentricity [little attention paid to probable meaning of others' actions].

The qualifications and experience of the expert(s) whose expertise is incorporated within the knowledge base is important. It is difficult to say who an expert is. For some tasks it may take up to twenty years of professional experience and knowledge to become an expert; whereas, in other tasks, the task might be so specific and unique that someone with a few months of experience may be called an expert. The expert is an individual, acknowledged by his/her peers, as being an expert. He/she generally has a keen acumen and an unusual talent for getting to the heart of the problem and solving it. The expert has typically built up a number of years of professional experience in performing the task, and has developed "rules of thumb" from experiential learning over the years in solving the task (4).

Acquisition/extraction of the expert knowledge is a major human factors concern. Knowledge acquisition is an iterative process in which many meetings with the expert are needed to gather all of the relevant and necessary information for the knowledge base. Because an expert system is only as good as its knowledge base, the collection of knowledge is critical for successful implementation and operation of expert systems.

Knowledge acquisition is perhaps the biggest bottleneck in expert system development. This is due to a number of reasons. First, the knowledge engineer must be familiar with the problem domain and specific task before he/she starts the knowledge acquisition sessions with the expert. A second major problem is the ability of the knowledge engineer to probe the expert's mind to obtain the pertinent facts and rules of thumb from the expert. The third is that biases are unintentionally imparted during the knowledge acquisition process by both the expert and the knowledge engineer. These biases inhibit the transfer of knowledge between the two individuals. One of the biases deals with intuitive statistical analysis (i.e., humans do not function well as intuitive statisticians). Another is the judgmental heuristic called "availability"; biases result due to the retrievability of instances. That is, when the size of a class is judged by the availability of its instances, a class whose instances are easily retrieved will appear more numerous than a class of equal frequency whose instances are less retrievable. Biases of imaginability and illusory correlation also play important roles in affecting an expert's judgement. Another bias relates to anchoring and adjustment (i.e., humans have a tendency to make judgements by establishing an anchor point and then making adjustments from this point). Two final biases are recency [humans are influenced more by recent events than by past ones] and concreteness [humans tend to use the available information only in the form in which it is displayed] (5, 6).

Humans are also susceptible to other errors and inadequate models which may influence the knowledge acquisition process (7). They include: (a) suboptimal level of schema abstraction, (b) sheer size/complexity of the schema, (c) inappropriate cues, (d) forgetting heuristics, (d) too little/too much information, (e) false recoveries, and (f) inappropriateness of certain verification processes.

There are five major ways to represent knowledge in the knowledge base-predicate calculus, production or inference rules, frames, scripts, and semantic or associative networks. In deciding among knowledge representation methods to incorporate into the expert system, a good rule of thumb is to select the approach that seems most natural to the expert. In other words, the knowledge should be represented in the expert system in the same manner that the expert is using knowledge when explaining a domain or task to the knowledge engineer (4).

As far as the nuclear utilities are concerned, the most important issues impeding the implementation of expert systems in electric power plants are the nature and quantity of verification and validation (V&V) which might be required by the NRC. In conventional software, V&V have well-established meanings. Verification is a determination that the software has been developed in a formally correct manner and in accordance with a specified software engineering methodology. Validation means demonstrating that the completed program performs the functions in the requirements specification and is usable for the intended purposes.

Present standards appear to be adequate for preparation of the inference engine, but, since the expert system goes beyond the procedures for conventional software engineering, the modularized, top-down, hierarchically decomposed design that makes conventional V&V possible is not applicable to the knowledge base. Also current V&V methods, which usually involve exhaustive testing, are generally considered inadequate for all but the simplest expert systems because expert systems - especially those operating under uncertainty or with incomplete data- have too many states to make exhaustive testing feasible. New approaches to V&V are therefore needed for expert systems. EPRI has an on-going research program (8, 9) which is aimed at satisfying the need. The program is oriented towards the development of a methodology for validating and verifying expert systems for nuclear power plant applications.

When an appropriate expert system V&V process is finally developed, it should be carried out by a group completely independent of the group(s) that designed and developed the expert system. In addition, the users should be represented in this V&V group. Expert systems V&V is related so intimately to the design that true independence may be difficult, but will be absolutely essential. The independence of the group that does V&V should be ensured by quality assurance procedures and organizational policy.

Human-Expert System Interface

The human-expert system interface is used to perform data collection, editing functions, and consultations. This interface almost always exists in an English-like format and includes a natural language that permits presentation of the expert system knowledge and processor explanations. Most expert systems have a degree of self-awareness or self-knowledge that allow them to reason about their own operation and to display inference chains and traces of the rationale behind their results.

The information that is presented to the human from the expert system via a computer-generated display (CGD) should be simple, clear, and understandable/ comprehensible. By understandability/comprehensibility, it is meant that the structure, format, and content of the display dialogue must result in meaningful communication. In other words, the "messages" displayed by the CGD must be interpretable by users, and the messages which they want to transmit back to the expert system must be expressible. During the expert system design process, the terminology, abbreviations, formats, and so on should all be standardized. The format should be familiar to humans and be related to the tasks they are required to perform with the information. The screen displays should be arranged so that the expert system users are not required to remember information from one screen for use on another (10).

Research on the understandability and compatibility of the expert system interface should be initiated. The reasons for this are as follows. The physical presentations to humans should consist of concise, high level information to support their cognitive functions. The nature of the display presentations to the users and the responses expected from them must be compatible with human input-output abilities and limitations (i.e., sensory, perceptual, and cognitive capabilities, human physical characteristics, and human physiological characteristics and capabilities). Succinctly, regardless of the overall expert system objectives, users have to be able to read the displays, reach the touch panel, and so forth. Otherwise there is a risk that the expert system will be inherently useless (11).

The design of the expert system interface should support effective use. A system is effective only to the extent that it supports the human (or crew) in a manner that leads to improved performance, results in a difficult task being less difficult, or enables accomplishment of a task that could not otherwise be accomplished. NRC staff members who were surveyed stated that design criteria should be established and followed. They suggested a program of research with the purpose of investigating the type of information and explanations that should be presented, the most appropriate presentation modes (i.e., text, graphics), and the frequency and content of the presentation of the information and/or feedback.

Does the information display support the way in which the user processes information, or is it merely determined by the way the software engineer describes the parameters of the system? The expert system information display must mesh well with the perspectives used by the human and the way in which the information is displayed should correspond to the user's mental model of the plant. People's view of the world, of themselves, of their capabilities, and the tasks they are asked to perform, or topics they are asked to learn, depend heavily on the conceptualizations that they bring to the task. In interacting with the environment, with others, and with the artifacts of technology, people form internal mental models of themselves and of things with which they are interacting (12).

One of the primary and most valuable features of expert systems is their ability to provide an explanation of the reasoning process used to solve a particular problem. These abilities are usually referred to as the explanation facilities. The features are very important because they enable the human to monitor the expert system's activities, understand why a conclusion was reached, and detect when the expert system has made an inference error. The human can take advantage of the explanation facilities to request: a complete trace for a consultation, an explanation of how a specific goal or sub-goal was inferred, or an explanation on why a particular piece of information is needed. However, the design of the explanation capability raises many human factors concerns. They include: what kind of explanation facilities should be included in the expert system (the user should be able to understand the expert system's behavior); should the explanation be presented as a trace of the rules that were considered by the expert system; should the expert system dictate an answer, or should it simply advise the human; what expert system information should be presented to the user and how should it be displayed; and should only the final conclusions be displayed, or should intermediate inferences be presented so that the user can understand and critique the expert system's performance?

"User friendliness" should also be considered in the design of the human-expert system interface. This is a "motherhood and apple pie" statement and a rather vague notion to implement. Some help is, however, available (13). Five criteria with which to base and measure user friendliness have been defined. They include: time for the human to learn, the speed of his/her performance with the displays, rate of user errors, subjective satisfaction of the displays, and human retention over time.

A number of other human factors concerns in regards to the expert system CGDs are: what should be the mode of interaction (i.e., graphics, alphanumerics, textual information, and/or mimics) between the operator and the expert system; is a textual display sufficient, or should graphics be added to enhance the human's comprehension; would a graphical presentation of the logic structure be helpful in understanding the conclusions reached by the expert system; is color coding required to call attention to certain parameters; how much control should the user have over the expert system; and should the expert perform any of its functions autonomously?

<u>Organizational Support</u>

The operator's ability to deal with an abnormal event or emergency, even at the level of reading information from the expert system, can be affected by the management style and the organizational support for the use of expert systems in the control room, as much as by the design of the information displays themselves. The ability of operators to respond to off-normal events is also affected by both fatigue and motivation. The structure and organization of shift work will affect operator efficiency due to disruptions in his/her biological circadian rhythms. A utility management, insensitive to comments by users about their working conditions and to suggestions in regards to expert systems, may obtain obedience to rules, but will not encourage participation in the pursuit of excellence. Civilians do not adopt dictatorial styles voluntarily and may resent them if imposed by management. Management practices are responsible, directly or indirectly, for establishing and maintaining an organizational culture that reinforces safety and the quality of performance. The formal structure, procedures, and practices of an organization bind the behavior of its employees and strongly affect the norms and perspectives they have regarding critical activities (11).

The design of many expert systems seem to be doomed to failure because managers/ engineers are more interested in designing the expert system than in first assessing the needs of the anticipated users. There is always a danger in beginning any design program without a complete assessment of the human needs. Machinists do not choose their tools before they examine their jobs; builders do not order their materials or plan their schedules until they have their blueprints. Why then, should engineers design expert systems without first specifying what the needs of the user are? A needs assessment of the user should be conducted prior to the design of any expert system so that the utility does not spend its money unwisely. During the needs assessment, needs and desires of the potential users should be identified and areas where an expert system could improve performance should be determined. The needs assessment should consist of three analyses, organizational, task, and person (14).

A function allocation and a division of labor between the human and the expert system should be conducted after the needs assessment, but before the system is designed. The anticipated user should be consulted during this process. The human should only be assigned those functions which he/she is most capable of performing and which best utilize his/her skills, knowledges, and abilities. In the past, allocation of functions was based on catalogs of "things computers do better" and "things people do better". With the current rate of technological development, however, existing catalogs are becoming obsolete, and this distinction may soon cease to be relevant in most situations. As expert system technology develops, the idea of fixed allocation is no longer appropriate. ORNL (15) outlined an approach to functional allocation that correctly emphasizes an iterative approach to the solution for conventional systems, but for expert systems, a different conceptual framework is required. The relation of the user to the expert system should be symbiotic. Human-related problems are symptoms, not causes, of underlying problems in the socio-technical system. Research should be designed to examine better methods and criteria for allocating functions between the human and the expert system. Research should also be conducted on how to design the expert system so that the human and expert system can support each other, request and give help as needed, and produce the most effective joint outcome.

The anticipated users of the expert system should be consulted during the entire life-cycle of the expert system so that they feel/believe that they are part of the design process. The users should be especially involved during the needs assessment, development, evaluation, and integration phases. Besides the users, engineers, managers, trainers/instructors, and human factors personnel should also work together during the design process so that there is cohesiveness between these types of personnel. When the expert system is introduced/implemented within the electric power plant, it should be thoroughly integrated with the other hardware, software, and tools in the user's work environment. The expert system needs to be introduced in a way which supports user acceptance. The impact of the expert system upon the other functions and tasks that the human performs should be evaluated and investigated.

Guidelines for the design, test, and evaluation of CGDs should be consulted during each expert system's life-cycle (10, 16). Human factors guidelines should also be utilized during the development of the expert system interface (17, 18, 19). There is some doubt, however, as to whether any of the existing guidelines are applicable to expert systems. The adequacy and applicability of the guidelines need to be investigated.

Related Topics

A potential safety concern is operator training. It may be necessary to evaluate the training program for any expert system that provides safety-related information or is involved in a nuclear plant safety system. Futhermore, a number of NRC staff members surveyed expressed concern that special training should be provided before the expert system is implemented in the work environment. They noted that the utility's training department should receive information and support from the expert system designers to the maximum extent.

The training program development for the expert system should begin early in the system's life cycle. Development should flow in unison with the design of software if at all possible. Anticipated users should also be involved during the preparation of the training courseware. Training materials developed for the expert system should be integrated with the existing user's training program. Features of the expert system should be discussed routinely during other systems training in order to show system interrelationships. The use of the expert system during normal/off-normal operations should be encouraged during training. Implementation of the training should take place via classroom, part-task training devices, and a full-scope simulator.

The expert system should not "overload" the users more than they already are; rather, it should simplify the required user tasks and unload humans of their mundane, routine, and tedious tasks. If at all possible, the expert system should reduce/relieve some of the existing workload, both physical and cognitive, on the user. Physical workload is defined as energy actually expended by the human; cognitive workload is defined as information processing which the user performs (20). Two questions which need to be asked any time a new expert system is introduced into the user's work areas are: does the system lighten or increase the human's physical workload; and does it lighten or increase his/her cognitive workload?

What humans will do under stress must also be considered. Will they be motivated/able to maintain their expertise when they have access to a powerful and intelligent assistant? Will they cease to consider themselves responsible for safety? Will they be able to detect when the expert system begins to provide incorrect answers, and to effectively resume control of the situation?

An evaluation of the effects of the expert system upon human performance (e.g., errors and time) should be conducted before it is implemented within the work environment. This evaluation is a post-audit to see if the expert system meets the objectives for which it was developed (i.e., making the user's job more effective and efficient). It should also be oriented towards making sure that the expert system does not confuse the user. Currently no method or tool exists with which to perform the evaluation, measure the performance of the expert system, and the effect of the system on human performance. New tools are, therefore, needed; they must have objective criteria that are quantitative in nature.

Research should be performed on the ways in which expert systems can assist human performance. People use data about the world in order to solve problems in that world. To do this, problem solvers must collect and integrate available data in order to characterize the state of the world, to identify disturbances and faults, and to plan responses. A basic fact in cognitive science is that the representation of the world provided to problem solvers can affect their problem-solving performance (21). Thus questions about expert systems can be reinterpreted to be questions about how they vary in their effect on the problem solver's information-processing activities and problem-solving performance.

A potential safety concern is the users' reactions to the expert system. Will they like the system and accept it? Will they be comfortable with an expert system and use it when needed? Will they believe that the system will work and that it is useful? Above all, will they trust and have confidence in the information presented by the expert system? Another concern is the possibility of over-dependence upon the expert system's guidance; a number of NRC staff members who were surveyed insisted that the user of an expert system may become too dependent upon its guidance, especially during off-normal events. They believe that an undue or blind reliance is liable to happen/occur. The expert system needs to be viewed strictly as a job aid or tool and should be used as only one of many inputs upon which to base decisions. It should simply advise the user, not dictate the course of action.

There is little understanding, at present, of what makes a person trust or distrust an expert system, the advice it gives, or the action it takes, and there is only the beginning of an understanding of the nature of the human cognitive processes that underlie the acquisition and assessment of evidence and the genesis of decisions on which trust is based. Yet these processes lie at the core of human control of expert systems and center on the nature of the user's mental models of the system, through which the user interprets the demands of the task. The National Research Council (11) stated that there is a need for laboratory-based facilities to evaluate human operator responses and acceptance of new technologies in artificial intelligence and expert systems.

FUTURE RESEARCH

Human factors issues related to expert system design and implementation have been identified. These issues will need to be studied further and evaluated thoroughly. A number of research programs will probably need to be initiated-some by the NRC, others by the EPRI, and a few by the electric utilities themselves. This research should be directed towards investigating concerns and answering the human factors questions.

NOTES

The research described in this paper was sponsored by the NRC under U.S. Department of Energy (DOE) interagency agreement 1886-8085-2B with Martin Marietta Energy Systems, Incorporated under contract number DE-AC05-84OR21400 with the DOE. The views and opinions are those of the authors and should not be interpreted or construed as the official position of the NRC.

REFERENCES

1. Expert-EASE Systems, Inc. Notebook of the Expert Systems Applications in Power Plants Seminar. Palo Alto, CA: EPRI, May 1987.

2. R. O. Duda and J. G. Gaschnig. Knowledge-Based Expert Systems Coming of Age. Byte, September 1981.

3. F. Hayes-Roth, P. Klahr, and D. J. Mostow. Knowledge Acquisition, Knowledge Programming, and Knowledge Refinement. Santa Monica, CA: Rand Corporation, R-2540-NSF, 1980.

4. J. Liebowitz. Introduction to Expert Systems. Santa Cruz, CA: Mitchell Publishing, Inc., 1988.

5. G. B. Davis. Strategies for Determining Information Requirements. International Business Machines Systems Journal, 21, January 1982.

6. A. Tversky and D. Kahneman. Judgment Under Uncertainty: Heuristics and Biases. Science, September 1974.

7. B. G. Silverman. Analogy in Systems Management: A Theoretical Inquiry. Institute of Electrical and Electronics Engineers Transactions on Systems, Man, and Cybernetics, SMC-13, 1983.

8. D. B. Kirk and A. E. Murray. Verification and Validation of Expert Systems for Nuclear Power Plant Applications. Palo Alto, CA: EPRI, NP-5978, August 1988.

9. E. H. Groundwater, M. L. Donnell, and M. A. Archer. Approaches to the Verification and Validation of Expert Systems for Nuclear Power Plants. Palo Alto, CA: EPRI, NP-5236, July 1987.

10. S. L. Smith and J. N. Mosier. Design Guidelines for the User Interface for Computer-Based Information Systems. Bedford, MA: The Mitre Corporation, 1984.

11. National Research Council. Human Factors Research and Nuclear Safety. Washington, DC: National Academy Press, 1988.

12. D. A. Norman. Some Observations on Mental Models. In D. Gentner and A. L. Stevens (Eds.), Mental Models (pp 7-14). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers, 1983.

13. B. Shneiderman. Designing the User Interface: Strategies for Effective Human-Computer Interface. Reading, PA: Addison-Wesley Publishing Company, 1987.

14. I. L. Goldstein. Training in Organizations: Needs Assessment, Development, and Evaluation (2nd Ed.). Monterey, CA: Brooks/Cole Publishing Company, 1986.

15. R. Pulliam, H. E. Price, J. Bongarra, C. R. Sawyer, and R. A. Kisner. A Methodology for Allocating Nuclear Power Plant Control Functions to Human or Automatic Control. Oak Ridge, TN: ORNL, TM-8781, August 1983.

16. P. R. Frey and W. H. Sides. Computer-Generated Display System Guidelines: Display Design. Palo Alto, CA: EPRI, NP-3701, Volume 1, September 1984.

17. Essex Corporation. Human Factors Guide for Nuclear Power Plant Control Room Development. Palo Alto, CA: EPRI, NP-3659, August 1984.

18. Honeywell, Inc. Human Engineering Guide for Enhancing Nuclear Control Rooms. Palo Alto, CA: EPRI, NP-2411, May 1982.

19. U.S. NRC. Guidelines for Control Room Reviews. Washington, DC: NUREG-0700, 1981.

20. T. B. Sheridan and H. G. Strassen. Definitions, Models, and Measures of Human Workload. In N. Moray (Ed.), Mental Workload: Its Theory and Measurement (pp 219-233). New York, NY: Plenum Press, 1977.

21. J. Rasmussen. Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering. New York, NY: North-Holland, 1986.

Supporting Users in the Field: Multimedia Delivery Vehicle for Expert Systems

RONALD E. JOY, BRIAN A. ISLE, and CHARLES P. BLOOM
Honeywell Inc.
Artificial Intelligence Department
Corporate Systems Development Division
1000 Boone Avenue North
Golden Valley, Minnesota 55427, USA

GEORGE H. QUENTIN
Electric Power Research Institute
Generation and Storage Division
3412 Hillview Avenue
Palo Alto, California 94304, USA

ABSTRACT

Expert systems (often referred to as knowledge-based systems) are rapidly moving from the research and development labs to field deployment. The success of getting these systems deployed and accepted in the field will depend on understanding and overcoming many constraints and problems of the potential user. Some of these constraints and problems are: the system must be usable in the required work environment; it must be easily accessible; and most importantly, the interface between the system and the user must be easy to use. If these constraints and problems are not understood and overcome, the system may be deployed to the field but it will not be used. In a paper presented at the EPRI Power Plant Control Conference in February 1989, Richard Shirley explained the criticality of the expert system user interface by saying:

> The user interface for an expert system is more than a display and an input device. Underneath the hardware is the software that makes the interface function for the application. It is the hardware and software together that determine the ease-of-use for the user. A poorly designed human interface will sink the expert system; it simply will not be used.

This paper describes part of the results of a research project undertaken by Honeywell for the Electric Power Research Institute. Specifically, this paper covers the project objectives to design, build, field test and deliver a general-purpose, multimedia, portable expert system delivery vehicle that includes both the user interface and the expert system in one package. The SA•VANT™ delivery vehicle meets the constraints and solves the problems mentioned above.

INTRODUCTION

The overall effectiveness of any expert system is a function of the knowledge applied to its problem-solving task and the delivery of that knowledge to the user. There is a direct relationship between how often an expert system is used and the functionality of the user interface. Often in gas turbine troubleshooting and maintenance applications, it is necessary to have access to documents such as schematics, electrical wiring diagrams, equipment block diagrams, and pictures of actual components themselves. Because these can be essential sources of information for a diagnostician, they should be included in an implementation designed to assist the user. In addition, the user's mode of interaction with the system will vary depending on the maintenance or troubleshooting application. Can the user interact with the system via a keyboard, or is voice input necessary? Can the user read a display, or is voice output necessary? If an appropriate mode of interaction is not available, the system will not be used.

The SA•VANT system, built by Honeywell for the Electric Power Research Institute (EPRI), is a portable and rugged multimedia delivery system for PC-based expert systems. SA•VANT supports input from both manual keyboard and voice recognition and provides output as text, speech, interactive video with graphics overlays, and printed hard copy. The SA•VANT design philosophy called for the implementation to be robust and versatile so that it could support a wide variety of expert systems, and modular so that its component parts and software could be upgraded easily to maintain it as state of the art.

USER NEEDS

If a delivery vehicle for expert systems is to be used in the field, it must meet the users' needs. In general, for maintenance and troubleshooting applications, the following user needs should be met: (1) It should be usable at a remote location; (2) the interface between the user and the expert system must be easy to understand; and (3) the system should be easy to use with minimal training. For the system evaluated in the field test described in this paper, the delivery vehicle met the following additional user needs: (1) One person must be able to carry it to the job site; (2) while it should be optimized for use by a standard two-person maintenance crew, it should also be usable by a single maintenance technician; and (3) the user should have the capability to use different media for both presentation and input of information.

SYSTEM DESIGN

In addition to the obvious design requirements of keeping it as small and as lightweight as possible, SA•VANT was designed to be fault tolerant, versatile and modular. It was designed to be fault tolerant so that it could detect its own equipment failures and isolate them with little degradation in operation of the expert system. It was designed to be versatile so that it could support a variety of expert system applications. Modularity was achieved in the design of the core software and hardware configuration, which will facilitate improvements to the system as the technology improves. The core software was designed to be easily integrated with future or existing PC-based expert system applications.

DELIVERY VEHICLE

The SA•VANT delivery vehicle has hardware and software components. It was designed to be lightweight and small enough to be carried by one user to the work site, where it is plugged into a 120-volt AC power outlet. No other connections are needed because SA•VANT contains the expert system, the user interface and the data storage.

Hardware Configuration

The present hardware configuration of the SA•VANT system is shown in Figure 1. It contains an 80286-based host computer, an 800-megabyte optical WORM (Write Once Read Many) drive, a custom expansion chassis with six slots, a printer, two flat panel screens and a custom keypad. It is the first prototype and is not yet optimized for efficient packaging. It is 23 x 18 x 6 inches and weighs approximately 40 pounds. A photograph of the prototype is shown in Figure 2.

A Grid computer is used as the 80286-based host computer with a Seiko 80-column printer attached to its parallel port. The Grid computer contains 2 megabytes of random access memory and a 40-megabyte hard disk. Attached to the Grid is a six-slot custom expansion chassis where add-on boards can be attached. Currently the slots are filled as follows: (1) Speech production board, (2) voice recognition board, (3) WORM controller board, (4) video production board, and (5) and (6) will be used for future enhancements.

The video images are displayed on a Hycom 7-inch diagonal, electroluminescent, flat-panel screen with 16-level gray-scale ability. The Grid has a 13-inch diagonal, plasma, flat-panel screen. The main keyboard has been replaced with a membrane keypad with a minimum number of larger keys removing the need for QWERTY typing abilities. The enlarged keys allow operation with bulky gloves for cases where gloves are necessary, such as electrical work.

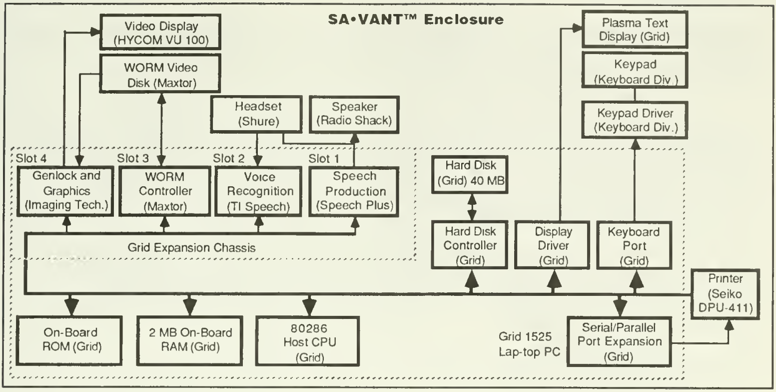


Figure 1. Current SA•VANT Hardware Configuration

Figure 2. SA•VANT Prototype

Software Structure

The two parts of the SA•VANT software structure are the expert system application software and the SA•VANT core software. While the expert system application software is very important, it is not discussed in this paper. The core software controls the access to the various media of SA•VANT by providing a well-defined device protocol that the expert system application follows.

The core software, written in ANSI Standard C, consists of a command dispatcher and several device drivers. This software is combined in a library that is linked with the expert system application software. The expert system passes commands to the core software dispatcher through subroutine calls. The dispatcher queues these commands, and upon request from the expert system, dispatches them to the device drivers. The modular design of the core software allows for easy replacement of the physical devices and the device drivers as new technology becomes available.

The core software acts as a buffer between the expert system application and the underlying hardware. It can detect and isolate a malfunction with a physical device, thus allowing little or no degradation in the execution of the expert system. Since the fault detection and isolation function also indicates what component (at the board or device level) is malfunctioning, repair of SA•VANT is reduced to the simple replacement of the indicated component.

FIELD TEST EVALUATION

Background

The SA•VANT system was developed to deliver expert systems to users in the field. The first expert system application developed with SA•VANT was for troubleshooting ground faults in GE MS7001E gas turbine control circuits in power plants. This was an excellent application for field test evaluation because the maintenance technician's tasks were characterized by interpretation of complex symptoms, isolation of logical faults and troubleshooting procedures that were often complicated. In addition, for this application there was a wide variability in the success rate and time to repair the control circuits based on a technician's expertise.

This was also an excellent opportunity for testing the SA•VANT delivery vehicle. The tasks performed by the technicians were often accomplished in cramped working quarters and required mobility among different work places. There was a wide range of environmental conditions such as extreme noise and poor lighting. The technicians used electronic test equipment, hand tools and printed documentation in these tasks.

The following steps were used in the evaluation: (1) The technicians were trained to use the new equipment; (2) ground faults were induced in the turbine control circuits; (3) the technicians were asked to diagnose the ground faults with and without the system; and (4) each of the technicians were debriefed after their session. Both the SA•VANT system and the expert system were evaluated.

SA•VANT and Expert System Evaluation

The evaluated areas of the SA•VANT system were the device hardware, the information presentation, the system operability and the user training. The device hardware evaluation was concerned with measurements of the physical operation, reliability and ruggedness of the system components. Included in the component evaluation were switches, microphone, speaker, video displays, computer and printer. The information presentation evaluation was concerned with the cognitive issues of comprehending the information presented by the system. Specifically, the understandability of the information presented, the quality of the guidance offered and the level and detail of the interaction/dialog with the user were evaluated. The evaluation of the system operation focused on issues of device portability, startup and shutdown, information readability, system timing, voice input and speech output. Finally, the user training evaluation was concerned with the ease of training-to-proficiency of the user on the expert system and the effectiveness of the user manual. The expert system was evaluated to determine if it could help both novice and expert technicians isolate ground faults without hindering either group.

Evaluation Results

The evaluation showed that the SA•VANT system and the expert system application were helpful. Some areas of the evaluation deserve special mention.

- All the subjects successfully isolated the grounded circuit in an average time of 25 minutes. The average time for the experts was 24.5 minutes and the average time for the novices was 26 minutes.
- Experts felt that using the system neither impaired nor slowed down their troubleshooting performance.
- Each subject received only one hour of training and practice using the system a few days prior to the evaluation. One could envision further time savings once an individual became more familiar with the system and its troubleshooting logic.
- Novices stated that without the system's help they would not have been able to isolate the grounded circuit.
- The text screen and printer exhibited no problems.
- The keyboard needed protection against multiple inputs, although the subjects found it easy to read and understand. Subjects who wore gloves had no glove-related problems with the keyboard.
- The video screen was too small and difficult for some subjects to read clearly.
- When using the speech output and not watching the screen, some of the subjects got confused. This confusion indicates the format of the speech output must be tailored to known limitations of the human information processing system.

The field test evaluation showed that SA•VANT could be used for more than its original purpose of delivering expert systems to the field. It can also be used as an intelligent document retrieval system and as an effective training tool. The expert system in the field test evaluation would retrieve and display schematics, drawings and pictures that pertained to the technician's work. Technicians who used SA•VANT in the field test stated that having timely access to the correct supporting documents enabled them to complete their tasks more efficiently. During field demonstrations, similar comments have been made by other technicians. Any application that is directed at this document retrieval capability could be developed for and delivered on SA•VANT.

It was evident that while using SA•VANT to diagnose actual equipment faults during the field test evaluation, the novice technicians were being taught an efficient troubleshooting strategy. They were able to learn from the expert system application because they could request an explanation for actions and a summary of the steps that were taken to reach a solution. SA•VANT could be used as a delivery vehicle for either computer-aided education or for a more sophisticated intelligent tutoring system. In either case, the combination of video images to show documents or physical locations, text description and intelligent student interaction would be a very powerful training tool. Furthermore, when learning about a task on a large machine, a student could take the SA•VANT tutor right to the machine.

FUTURE ENHANCEMENTS

SA•VANT was designed so that as new technologies become available, it would be easy to upgrade. Future enhancements include improvement in the video storage and presentation, improvement in the voice input and the speech output capabilities, a decrease in the size and weight, addition of data acquisition capabilities and improvement in the keyboard.

Video storage and presentation will be improved by decreasing the video frame display time. This will be accomplished in several ways. The host computer will be upgraded from an 80286 to an 80386 CPU. The

WORM drive will be replaced with one that uses a more sophisticated cache algorithm and faster data transfer rates. The digital video display board will be upgraded to include video compression/decompression algorithms that will give a 10:1 reduction in the data required to store a single frame of video. Presentation of the video information will be improved by increasing the video display's size from the present 7-inch diagonal to a 12-inch diagonal. Full motion video will replace the existing video system as soon as Digital Video Interactive (DVI) technology becomes available.

The voice input and speech output will be improved by incorporating the results of ongoing research on optimizing voice interaction between the user and SA•VANT by tormulating the data more closely to natural dialog.

Several methods ot decreasing the size and weight of SA•VANT are being investigated. These include switching to a larger single screen and utilizing a video window, and the adoption of more compact components such as a half-height WORM drive.

In the near future, SA•VANT will include a data acquisition capability to collect data from control systems or from auxiliary sensors. The data can be used to keep track of machine performance to predict impending failures or to provide enhanced diagnostics and troubleshooting capability. Initial work will be to provide data acquisition for vibration monitoring sensors and collection of on-line control data from Westinghouse gas turbines.

The improvement of the keyboard is now being done. The mounting platform is being stiffened and a new keypad and software to protect from multiple key presses is being developed.

CONCLUSIONS

The multimedia interface of SA•VANT makes it an effective and useful tool for the delivery of expert systems to the field. The authors believe that any PC-based DOS expert system can be easily ported to the SA•VANT delivery vehicle. Expert systems built using Prolog and tools from General Electric, Texas Instruments and Honeywell have been ported to SA•VANT. SA•VANT is easy to learn and use. With the appropriate knowledge base, it will allow inexperienced users to function as experts in limited domains. SA•VANT may also be used as a training tool for intelligent document retrieval and as a vehicle for delivering nonexpert system software.

Future refinements to the SA•VANT system include making it smaller and lighter, refining the voice input and the speech output, modifying the keypad and keystroke software, and adding a larger and higher resolution video screen. As Digital Video Interactive (DVI) technology becomes available, it will replace the existing video system, thus providing full motion video.

ACKNOWLEDGMENTS

The work described in this paper was performed under the EPRI research project (number RP2562-1) titled SA•VANT. The authors gratefully acknowledge the support provided by Arch Butler of Honeywell and Clark Dohner and Al Dolbec of the EPRI Generation and Storage Division.

Lessons in Deployment of Successful On-Line Expert Diagnostic Systems

K. ERIC HARPER, JAMES C. BELLOWS, and ROBERT L. OSBORNE
Westinghouse Electric Corporation
Orlando, Florida 32826-2399, USA

ABSTRACT

There is a need for on-line expert diagnostic systems in the utility industry. The goal of the systems should be to supplement existing procedures for handling operating and maintenance decisions, and duplicate the diagnoses and recommendations of the experts who design, service, and maintain the power plant equipment. For multiple installations where repeat diagnoses are infrequent, like utility power plants, a centralized system configuration is best. Other considerations are rulebase size, project funding, data management, data storage, knowledge documentation, end user, and graphic requirements. A centralized approach uses hardware and software locally at the plant sites and at a central support location. Staffing includes knowledge engineers, computer scientists, experts, and diagnostic operators. Careful planning and management of rulebase development and maintenance is important for success. The investment can payoff in reduced forced outage rates and increased availability of power plant equipment.

NEED FOR EXPERT DIAGNOSTIC SYSTEMS

There is a growing need for on-line expert diagnostic systems in the utility industry. On-line expert systems translate continuous sensor data into a description of the condition of the monitored equipment. Increased visibility of the present and future conditions of the power plant make it possible to lower operating costs. Equipment life can be extended and forced outages avoided by making informed decisions on how to run the plant. The savings are substantial, especially on a utility's largest, most efficient units.

Currently utilities obtain this visibility with large monitor systems measuring thousands of variables critical to the proper operation and protection of the plant. These systems are designed to alarm if a variable exceeds one or more limits, allow the operator to trend one or more variables, and display the values superimposed on diagrams of the equipment to facilitate operator identification of the physical location of the variables.

Although these systems are useful in data presentation and manipulation, what the operator needs is:

- Minute by minute status of the power plant,
- Specific recommendations if and when action is required,
- Prioritization of the actions so that the most critical situations are clearly identified,
- Potential consequences if action is not taken.

This help is even more critical during high activity periods like startups or other plant transients when the number of variables in alarm is large, variables are changing rapidly, and the time to assess each situation is limited.

On-line expert diagnostic systems are available and are designed to address these operator needs. They have been in everyday control room use for over four years with total experience exceeding thirty-five unit years. An indication of their effectiveness is shown in Figure 1. The figure traces availability and forced outage rate for seven large electric power generators from 1984, before on-line expert diagnostic systems were installed and operational, and from 1985 to 1988 when the seven systems have been operational. An average increase of seven days availability was obtained. Using $500K per day as the cost of unavailability, this translates to $3.5M per unit in savings each year.

The goal of on-line expert diagnostic systems should be to supplement existing procedures for handling operating and maintenance decisions. The system should duplicate the diagnoses and recommendations of the experts who design, service, and maintain power plant equipment. This paper is based on the experience gained in implementing and operating an effective on-line expert diagnostic system, and explores many of the challenges that should be addressed.

GenAID INSTALLED IN 7 UNITS OF ONE UTILITY

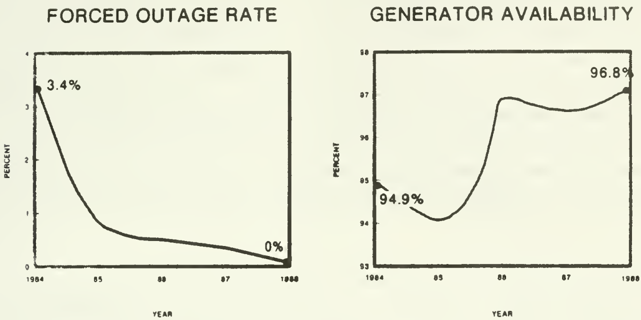


FIGURE 1

PHILOSOPHY AND SYSTEM REQUIREMENTS

The success of the generator on-line expert diagnostic system is due to several factors. First, a centralized approach is used to track and satisfy the needs of the utility customer, giving them access to a large base of turbine generator expert knowledge. This design makes it possible to control the changes made to the rulebases, reduce the computer resources necessary to support the power plants through operating transients, and provide the capacity to hold the thousands of rules necessary to deliver complete diagnosis of the generator. Second, the on-line diagnostics service business is set up with access to a continuous cash flow through other corporate resources to support the long term investment needed to deliver quality and comprehensive scope diagnostics. Last, the expert system is supported by human diagnostic operators and technical assistance.

To achieve the same success a requirements specification should be written identifying the system's users, components, and environment prior to the purchase of either software or hardware. These requirements have a direct effect on the size and type of hardware and software that needs to be purchased or developed.

Centralized Design

Knowledge can reside in the power plant or be located remotely. For multiple processes where individual installations have infrequent repeat diagnoses, like utility power plants, a centralized configuration is best. The advantages of a central location for all diagnostic knowledge bases include:

- o Staff for the varied skills necessary for knowledge base development and maintenance is in one location,
- o Knowledge gained from one plant can be quickly applied to all connected plants,
- o System cost is reduced by data filtering and sharing the large computer capacity required during individual plant high activity periods such as startup and other transients.

Systems which are sophisticated enough to maintain the operator's confidence in the diagnoses contain thousands of rules and diagnose hundreds of conditions on critical equipment such as the electric generator. If the knowledge and computer resources are located separately in each plant this investment must be duplicated for each site.

Diagnostic operators at the central location track the diagnoses and backup the plant on-screen notification of abnormal operating conditions. Transition to regular expert system use is eased by initially providing a human interface to the plant. These personnel provide additional support for operators and plant maintenance personnel. This is similar to the cost effectiveness achieved by a utility's central maintenance crew.

Rulebase Size

For a properly maintained rulebase, the size will increase over time. This is analogous to a human expert. As the expert gains more experience, his knowledge increases and thus the quality of his work can be enhanced over time. For an electric power generator, the diagnostics presently identify over 500 conditions and utilize rulebases with 3000 to 4000 rules. Initially they were half this size.

Continuous Cash Flow

Expert systems, like the humans they emulate, grow and change with exposure to new data. Funds should be allocated each year to support the changes necessary for successful operation of on-line expert diagnostics.

Data Management

On-line expert diagnostics system load is affected by the volume of data received at the central location. A deadband method should be used to filter data transmissions from the plant site. Unless a variable changes by more than a pre-determined amount, it is considered constant. This strategy means that variables which change minimally under normal conditions are usually represented by few data points. If they become active, the number of transmissions can increase to provide an accurate trend. The reduction in average load can be a hundred fold. With the dead-banded data strategy the diagnostic computer should be sized to handle startups, typically a ten-to-one increase in data flow. This strategy can significantly reduce both the database load and the expert system load, since only significant changes are either saved or diagnosed.

Continuous Data Storage

All the data should be archived for the knowledge base maintainers to enhance the quality of diagnosis. Critical precursors of conditions can be missed if data is recorded only when an alarm occurs. The number of opportunities to learn from actual events, and thus increase diagnostic system quality, is limited due to high power plant reliability. The actual size of the database depends on its implementation, the number of connected plants, and the number of points transmitted to the central location.

For example, a potentially damaging condition in electric power generators is cracked conductor strands. If a large percentage of strands are cracked the conductor can arc, requiring subsequent repairs that can be as costly as a total winding replacement. In any given year only a few generators may have cracked strands. The trends related to predicting cracked strands are subtle and develop over a long period of time. If data is not taken continuously in advance of an alarm, the cracked strand incident will yield little usable information that can help prevent the next incident.

Knowledge Documentation

Documentation is critical to the quality of the diagnostic system, and crucial for efficient maintenance. When the number of rules grows into the thousands, the time to determine a knowledge base problem, identify a solution, and verify that the identified changes will not adversely affect other areas of the knowledge base becomes very expensive in engineering time without good, usable on-line documentation which is always up to date. The expert system shell should have a document facility which allows unlimited text entry. Constructed in this manner, the documentation is generated at the same time the rulebase is developed or modified, and it is up to date.

End User

Choice of the end user has a significant effect on the ultimate size and value of the system. A knowledge engineer user generally has the capability and interest to recognize diagnostic quirks or perplexing output, and compensate for them by interpreting the output. This type of user can live with a smaller, less sophisticated system. On the other hand, if the system is to be used by a number of plant operators 24 hours a day when immediate expert human diagnostic help is not available, then the system should be large and sophisticated to provide sufficient on-going accuracy to maintain operator confidence. Without this confidence the operator will stop using the system in everyday practice and the entire investment is lost.

Graphics

Graphics are very important in the presentation of information to the operator. A minimum of knowledge and effort should be required to operate the graphic interface. The display should locate each active condition on plant equipment diagrams.

DETAILED REQUIREMENTS AND RESOURCES

On-line diagnostics is the process of converting automatically collected data into information that can be used by a plant operator to make informed decisions in less time. Typically the equipment required is for data acquisition, communications, CPU resources, data/results display, and data storage and retrieval. These components are purchased and installed once as an initial expense. However, on-line diagnostics has been a continuous effort in terms of maintaining and enhancing the knowledge base, and enhancing the process itself. For that reason a staff is required to support the on-line diagnostics operation during the life of the system. With the centralized diagnostics philosophy, the hardware components required for on-line diagnostics are located both at the plant sites and in a central location relative to the plants. The installations are connected via a data network that allows information transfer and other remote access. The software programs required for on-line diagnostics run on computers located at the plant sites and in a separate central location. The programs transfer information via process-to-process communications over a network. These requirements are addressed by purchasing or developing software programs.

PLANT BASED REQUIREMENTS

Plant Data Center

Hardware. On-line diagnosis is driven by automatic data input. Data for a plant process is usually available as part of the monitor and control equipment provided by the manufacturer. Often additional points may need to be added to produce diagnoses of acceptable quality. Data scan times and resolutions should be consistent with the time constants and signal levels of the plant process in order to determine trends and capture transient events. If significant additional measurements are required, it may be more cost-effective to install a state-of-the-art data acquisition system rather than expand existing capability.

Software. Commercially available data acquisition systems provide different levels of features when it comes to filtering, engineering units conversion, secondary variable calculations, etc. Software running at the plant site should deliver validated point values that can be entered directly into the expert system. Most modeling and state estimation should also be performed at the plant site due primarily to the large amount of data which would otherwise need to be transmitted.

Plant Database

Hardware. Computer disk and RAM memory resources are needed to maintain short-term records of acquired data at the plant site. This is necessary to calculate secondary variables based on slopes and averages, which are then used by the expert system in the diagnosis. The database also supports plant display trending and analysis.

Software. Maintaining a database at the plant site provides storage for sensor and calculated variable point histories. The histories are implemented as ring buffers where new values replace the oldest values. All recent data points transmitted to the expert system should be saved as a side effect of the transmission. The newest value for each point is made available to secondary variable calculations to implement running averages, slopes, and state change detection. Point values should be displayed locally in data lists, trends, or crossplots.

Plant Display

Hardware. The operator needs a graphic display which is oriented towards diagnostics to integrate this function with the normal duties of monitoring and controlling the plant process. This requirement can be satisfied with an additional graphics terminal in the plant control room or where possible, display information can be integrated into existing control room displays.

Software. The plant displays should be oriented towards diagnostics. In other words, the primary information is what condition is beginning to develop, and secondary information is the data to support the diagnosis. Operation of the displays should be intuitive or easy to learn because the audience is for the most part plant operators with many other responsibilities and little familiarity with computers.

CENTRAL LOCATION BASED REQUIREMENTS

Communications

Hardware. Communications is necessary to transfer information between the various computers and displays due to the distributed nature of on-line diagnostics. Recognizing that critical lines of communications can be affected by circumstances outside the plant's control, backups should be included to maximize reliability. The transfer bandwidth should be sufficient to handle both steady-state conditions and the large loads associated with plant startups and shutdowns. A wide area network maintained as a corporate resource can have an availability of over 99 percent.

Software. Data transfer between plant sites and the centralized expert system should be able to survive intermittent network malfunctions without loss of data. Data acquisition at the plant still continues if the link is lost, storing the information for later forwarding when the link returns. Similarly, pending diagnoses and recommendations coming from the central site should be stored and forwarded when the link returns. Although loss of communications delays the data and associated diagnoses, the information still has value and maintains continuity in the databases.

Expert System

Hardware. The heart of on-line diagnostics is the expert system. Sufficient CPU, memory, and disk resource is needed to:

- Deliver diagnoses and recommendations in a timely manner,
- Handle large numbers of rule firings triggered by transient data,
- Maintain active knowledge in memory for fast access,
- Provide on-line database access for expert system enhancement.

Typically a super-mini or mainframe computer is used for the expert system. It should be sized to handle the high capacity required for plant transients. The total investment is reduced for a centralized system because of transient data load leveling over many plants compared to having full capability at each plant.

Software. On-line diagnosis requires several different tools and features for creating, testing, and using rule-based knowledge. All the programs have the same inference engine and produce the same results, but the way the information is presented varies with each tool.

Knowledge Editor. An interactive editor is needed to capture knowledge. The editor should have a well-defined knowledge representation to reduce training and rulebase maintenance costs. It should be tailored to support the people who are responsible for making the expert system a success. This audience can be knowledge engineers, or better yet, the experts themselves. The editor interface should support casual users with menus, and sophisticated users with direct commands.

Entering knowledge into a rulebase is simplified by an editor which is basically "fill in the blanks." Module testing should be performed in the editor because developers want a good feeling that what they are coding is correct when entered into the computer. This ease of loading, editing, and testing allows the knowledge engineer to concentrate on the knowledge and can significantly reduce the time and effort to create a rulebase.

Verification. The second tool is used for verifying the rulebase with simulated plant data. Verification is the process of proving that the rulebase does what it was designed to do. The verification interface should provide detailed information about intermediate hypotheses and results, and present time-based diagnoses in terms of the sequence of events that lead to the conclusion. Verification is more productive and successful if all the information related to the test is available without having to switch screens or resort to hardcopy.

Production Diagnosis. The power of on-line diagnosis is that it automatically processes plant data. An environment is needed that once started, accepts new data from the network and produces a corresponding diagnosis. The environment should allow external access to view intermediate hypotheses for troubleshooting purposes. The crucial measure of production performance is the time delay between when the data is received and when the corresponding diagnosis goes out. The production environment should monitor and record this metric.

Central Database

Hardware. Crucial to the long term success of on-line diagnostics applications is the storage of acquired data for future analysis. The information is used to enhance the knowledge base as incidents occur and the characteristics are recorded. In addition, it is essential that the data which triggers diagnostic results be reproducible in order to verify new knowledge additions. In the plant, this level of data quality has not been available to results and design engineers in electronic form. Typically monitoring records are archived on paper logs or magnetic tape, making it difficult to import the information into analysis programs. A much deeper understanding of the plant equipment is realized when on-line data is available.

To fill this database requirement, sufficient disk resources are needed to maintain at least six month's worth of data on-line in a database. Magnetic tape or optical disk resources should be used to archive older data.

Software. Sensor data should be stored as a side effect of receiving points from the network at the central location. In this manner the central database duplicates the short term histories at the plant, and both diagnostic operators and plant operators see the same information. The database interface should make it easy to select and review information. Point values retrieved from the database should be in a form that can be directly entered into the expert system.

Diagnostic Operations Center

Hardware. On-line diagnostics is a partnership between the provider of the diagnostics service and the utility plant operator. For the partnership to work the plant operator should have the perception that the service will contribute to the plant's success. The diagnostic operations center is a twenty-four hour, seven day hotline to support the plant. Personnel in the operations center monitor all the plants on a twenty-four hour per day, seven days per week basis, and back up the in-plant diagnostic screens when abnormal conditions arise.

This requirement is fulfilled by a room with displays that duplicate and consolidate the individual plant diagnoses, along with electronic mail and voice communication to the plant control rooms.

Software. The central operations site consolidates the resources necessary to monitor and maintain diagnostics. The diagnostics operator needs to see everything that the plant operator sees to effectively communicate with the plant. For that reason the operations environment is a duplicate of the plant displays. The interface should allow access to each plant's data, diagnoses, and recommendations via menus and direct commands, and make it easy to log shift activities for customer reports.

Personnel

Knowledge Engineer. The role of the knowledge engineer has changed dramatically with on-line diagnostics. It used to be that the knowledge engineer was only responsible for interviewing experts and representing knowledge in terms an expert system could use. This scope was based on the assumption that input data is error-free and the knowledge engineer is the one viewing the diagnostic results. On-line diagnostics requires an expanded scope for the knowledge engineer. Their responsibility is ownership of the entire information process, from data to diagnosis, including:

- Data acquisition integrity and sensor validation
- Engineering units conversion
- Modeling and secondary variable calculations
- End user data presentation
- Knowledge acquisition, maintenance, and configuration control
- Knowledge documentation
- Knowledge verification and validation
- End user diagnostics and recommendations
- Feedback on system performance

This "end to end" responsibility is necessary because each of the above items can affect whether a diagnosis is correct or not, and whether an operator or user takes action based on the information provided him. If he takes no action then the diagnostic system will not produce savings for the utility.

To address these requirements the knowledge engineer needs a combination of technical and people skills. The technical skills are needed to understand the equipment, recognize abnormal operating conditions, and effectively use computer resources. The people skills are needed to form alliances with experts to enhance the quality of the knowledge and with other engineers and operators to maximize the effectiveness of the system.

To bring everything together the knowledge engineer should understand the tools used to create and maintain the knowledge. A successful approach has been to teach knowledge engineering to domain specialists, such as mechanical engineers. Domain knowledge is required to clearly structure the knowledge elicited from experts and to intelligently resolve conflicting expert opinion. An advanced degree is not required, but curiosity about how things work and a willingness to make decisions in the face of uncertainty are necessary. A requirement for success is that the knowledge engineer view himself as the champion for the project.

Computer Scientist. One of the advantages of expert systems is the separation of knowledge from the expert system shell. The knowledge engineer owns the knowledge. A parallel function is ownership of the expert system shell and associated on-line diagnostics processing. This responsibility requires the skills of a computer scientist. The synergy between the two functions produces an on-line diagnostic system that meets the needs of the plant. Close communication and cooperation are necessary for the partnership to be successful.

The computer scientist should create an environment that reduces the workload of the knowledge engineer, making him more efficient and productive. This environment includes:

- A knowledge representation that:
  - parallels the real world
  - models human thought processes
  - maps on-line sensor data into the knowledge
  - maps diagnoses and recommendations to the results display
  - allows hierarchical organization of the information
  - integrates documentation with the knowledge
  - can be presented graphically

- A knowledge editor that:
  - produces usable knowledge with minimum entry effort
  - prompts the novice user for information
  - allows direct access of functions by sophisticated users
  - encourages creation of documentation
  - integrates configuration control
  - checks and flags errors early in the development process
  - supports modular design and testing
- Test tools that:
  - verify fundamental knowledge design
  - support regression analysis of knowledge results
  - simulate incident scenarios
  - provide access to intermediate diagnostic results
  - reduce edit, test, debug cycle times
- An integrated system that:
  - reliably transforms data into diagnoses
  - measures and reports its own performance
  - is easily maintained and enhanced
  - provides guidance in the use of the system

To fill these requirements the computer scientist should have a combination of technical and people skills. The technical skills are needed to create and maintain software products in the monitoring and expert system domain. The people skills are needed to form alliances with knowledge engineers to identify when new experiences require system enhancements, and to enhance the quality of the expert system.

Experts. On-line diagnostics supplements and multiplies the diagnostic power of experts. The goal of the expert system is to duplicate the expertise of the people whose time is at a premium. Thus these experts can effectively be in more than one place at a time when their knowledge is utilized in an expert system. The expert is freed from routine problems and can then devote his time to new problems and to expanding the knowledge rather than conveying it to others. The expert is responsible for making sure that the knowledge is quantitatively accurate and logically consistent. In the end, the knowledge engineer actually becomes the expert for existing knowledge and the main archive for the information is in the knowledge base.

Diagnostic Operator. The diagnostic operator at the centralized site is like a shift supervisor in a manufacturing facility. His job is to make sure the process is operating smoothly and to expedite any situations that could interrupt production. The diagnostic operator should be backed up by technical people who can be contacted in the event of abnormal operating conditions. The diagnostic operator is responsible for reviewing all the plants' diagnoses and notifying the plant operators when problems arise. On-line diagnostics succeeds because of this personal contact, emphasizing a partnership between the diagnostics provider and plant consumer.

To fill these requirements the diagnostic operator should have a combination of technical and people skills. The technical skills are needed to understand the plant process to the degree of discriminating between normal and abnormal operation. The people skills are needed to form alliances with plant operators to influence the operation of the plants.

Diagnostic Knowledge

Knowledge acquisition is an evolutionary process. On-line diagnostic knowledge is the relationship between sensor readings and equipment condition. Without these relationships the diagnostic expert system will not be successful. This information can be acquired by experience, or from an understanding of the basic principles that govern equipment performance. A good place to start is with the manufacturer's installation, operation, and maintenance manuals. The next step is to consult with experts who have designed, operated, and maintained the equipment. Last, if the machinery has a monitoring system with a data archive, records can be reviewed for relationships.

RULEBASE DEVELOPMENT AND TESTING

A disciplined approach to rulebase creation is required if costs are to be contained. First the knowledge engineer, who is already skilled in the general domain, familiarizes himself with the system. He uses instruction manuals, general design manuals, and possibly one expert as a mentor to develop a qualitative understanding of the system to be diagnosed. When he finishes this phase, he writes a specification of the diagnosed conditions, associated recommendations, and what sensors or monitors will be required to diagnose each condition. This specification is reviewed by management and experts for appropriateness and technical feasibility.

After the specification is approved, the knowledge engineer interviews the experts and determines the details of the relevant ideas and their relationships. When this task has been completed, he codes, documents, and tests the rulebase. The fatal mistake in on-line diagnostics is to produce wrong or misleading diagnoses. Once user confidence is lost, it is extremely difficult to recover. For this reason a rulebase should be carefully tested before it is used. Testing is in four stages. The first is off-line test cases containing real or synthetic data. This stage is usually conducted along with loading and documentation to be sure that the various parts of the rulebase work as the knowledge engineer expects. The second test is an exhaustive evaluation of variables to determine that significant deviations before and up through alarm levels produce an appropriate diagnoses. The third test is end-to-end, where the rulebase is placed on-line and the sensor values are adjusted at the plant and the appropriate diagnoses are verified as present. Finally, actual on-line data is applied to the rulebase over a period of test time. During this phase, the knowledge engineer watches the diagnoses extremely carefully and may modify the rulebase to take into account subtleties that the experts had unconsciously glossed over during the interviewing process.

The next step in development of a rulebase is a design review. In this step, the final product is reviewed against the original specification for completeness. It is reviewed by experts for technical accuracy, and then released to the customer application. The last step is a continuing effort to expand and enhance the capability of the rulebase as new or enhanced knowledge becomes available. Like human experts, an expert system rulebase should become ever more knowledgeable if it is to remain valuable.

MAINTENANCE

Hardware

Experts and expert systems rely on the accuracy of data to draw correct conclusions. These conclusions should include diagnosis of both equipment malfunctions as well as instrumentation malfunctions. Well constructed expert systems are able to continue to operating effectively when monitors malfunction, but good sensor maintenance is required to make any diagnostic system work well, including systems where humans alone are required to make the diagnosis. For reliable sensors, such as thermocouples, this maintenance usually does not exceed annual calibration. For less reliable sensors, such as some of those that monitor plant chemistry, daily attention may be needed. Although the diagnostic system can enhance the efficiency of sensor maintenance, the power plant staff should still be present to do the maintenance.

The computer equipment also requires maintenance. I/O should be recalibrated periodically. Moving parts, such as disk drives wear out. Power supplies can be cut off. Chips malfunction. Each component requires technicians trained in its repair or service contracts with the manufacturer to be sure that it is on-line when it is needed. The service should be prompt, because the diagnostic system is unavailable if one of its major components breaks.

Software

On-line expert diagnostics software, like any other software product, goes through a process of revision. Each new release contains defect repairs and added features. With licensed software, the only maintenance necessary is to install and verify new versions, and report any problems to the vendor. Internally developed software requires a higher level of support. A good system of review and testing procedures should be implemented to reduce the number of non-conformances to software requirements, and to detect and filter out errors before general release of the programs.

Knowledge

Rulebases are continually being enhanced. Any time that the rulebase does not diagnose a significant condition, or diagnoses a condition erroneously, it should be carefully examined and modified. This modification usually adds rules to the system. Often it adds conditions as well. Another driving force for enhancement is the suggestion by a customer that a particular condition would be useful.

Data Base

The central database has a finite size and capacity for storing point values. Therefore it is necessary to periodically off-load older data from disk to magnetic tape. This maintenance activity should not interfere with normal production operation. If the data is needed later for analysis the values can be re-loaded from tape. As new applications are added the database should be configured to recognize new unit designations and point names. This can be automated to some degree.

COST ANALYSIS

Development and ongoing costs for on-line expert diagnostic systems fall into two basic categories: personnel and facilities.

Personnel Costs

Based on over five years experience the cost for each rule runs one to two man-hours of a knowledge engineer's time. In addition the time of a systems analyst and the equipment experts would add another one man-hour bringing the total engineering cost to two to three man-hours per rule. The cost referred to here is the total manpower cost for each verified rule which is actually providing information to the control room operator on a continuous basis. This would include the time spent to throughly understand the equipment, identify the sensors and the conditions to be diagnosed, a preliminary design review, interview the experts, design and write the rulebase, test the rulebase both off-line and on-line, a final design review, and a complete documentation package. It does not include development of new knowledge.

The number of rules required for each major component such as a generator will be in the area of two thousand rules initially and increasing to four thousand rules in several years. If the rulebase is much smaller than this, the equipment will likely not be covered thoroughly enough to insure the operator's confidence in and use of the system. Using common commercial rates the development cost will be up to one million dollars per component.

As long as the rulebase is in operation, at least one and preferably two engineers should maintain their knowledge of the details of every rulebase to a sufficient level that emergency maintenance and necessary enhancements can be made without excessive re-learning time. This appears to be possible in actual practice only by having such personnel actively working with the rulebase on a continuing basis.

Computer Costs

The developer should decide if diagnosis is to be done during startup, shutdown, and significant load changes or only during quasi-steady state conditions. The answer to this question is critical to computer sizing, especially where the system is to be located in the power plant and handle one unit. Our experience has shown that the computer load is more than an order of magnitude higher during startup than at a steady load. If a centralized approach is used this has significantly less effect on required computer capacity because only one or at the most two units would be starting at the same time.

Diagnostic systems where the number of rules is in the low hundreds can be handled by PC-sized computers. When the number of rules is in the upper hundreds or thousands the computer capacity must be in the multi- MIPS range with significant size RAM and hard disc storage capacities. Typically, this size of computer for a single unit would be in the $300K to $500K range. This would provide no backup computer capacity. In addition, service cost on this size machine would run approximately 10% of the purchase price per year. In addition, some computer technician or engineering effort would have to be available for program backups, restarts, and other on-going tasks. Thus the initial investment cost for an entire power plant will be in the millions with a significant percentage of this required each year for both software and hardware maintenance.

CONCLUSIONS

The use of on-line, expert system based diagnostics has shown to have a significant effect in reducing forced outage rates and increasing availability of power plant equipment. The resources, both human and financial, required to construct and maintain an effective diagnostic system are considerable. Years are required to develop a system which reliably provides on-line diagnostics to the control room operator. Utilities contemplating such diagnostic systems should carefully consider the total cost of in-house development versus the use of systems already available.

Application of PLEXSYS in Nuclear Power Plants: Technical Specifications Monitoring and Maintenance Management

SIAVASH HASHEMI and LAURA J. PATERSON
PLEXSYS Project
IntelliCorp
1975 El Camino Real, West
Mountain View, California 94040, USA

JOSEPH SOMSEL
Pacific Gas and Electric Co.
San Francisco, California 94106, USA

ROBERT E. COLLEY and RANDALL S. MAY
Electric Power Research Institute
3412 Hillview Avenue
Palo Alto, California 94304, USA

1. Introduction

The purpose of this paper is to describe a typical application problem and the development of a prototype expert system using PLEXSYS (1, 2) and KEE (3). The PLEXSYS model editor is used to build a basic domain model that represents the components and their connections. Structure is then added to the basic PLEXSYS model by defining additional units and slots for the KEE knowledge base and by adding rules using the KEE RuleSystem. Finally, an additional layer of structure, rules and customized user interface is added to complete the prototype expert system.

2. Background

An important class of maintenance planning problems involves the determination and evaluation of "tagout boundaries" for components scheduled to be temporarily removed from service for inspection or maintenance (4). The tagout boundary for a subject component is the minimum set of boundary components, such as valves or circuit breakers, that must be physically and/or administratively disabled to appropriately isolate the subject component from electrical and/or hydraulic systems. Administrative disabling is typically achieved by hanging on the control device, a warning tag that forbids changing the isolated component's state.

Constraints on component maintenance and tagouts are implied by the plant Technical Specifications (Tech Specs) and in particular the Limiting Conditions of Operation (LCO). The LCOs define the minimum set of system functions that must be active for a given operational state. The maintenance staff must ensure that no planned maintenance action will compromise these required functions. As the LCOs are quite complex, and maintenance must be performed simultaneously on a variety of components from different subsystems, confirmation that a maintenance plan is in conformance with Tech Specs may be a very difficult task.

In a typical nuclear power plant, maintenance planning activities are supported by access to relational data bases that describe the maintenance activities, plant components, relevant procedures and other essential information. For a general plant application, the range of possible situations and solutions is too broad for direct solution by a scheduling algorithm, and software tools are provided as aids to human planners who can make use of heuristic rules as well as their knowledge of the latest revisions to the plant systems and administrative requirements. Prior to a major outage, these efforts may involve dozens of human planners that must coordinate their efforts at each step. These characteristics make the tagout planning problem well-suited for an expert system approach, and rule-based representations of LCOs in a maintenance planning context have been previously published (5, 6).

The present paper describes a prototype expert system that uses a model-based reasoning approach to support maintenance planning and tagout decisions. The prototype described here has been implemented for the Residual Heat Removal (RHR) System for the Diablo Canyon Nuclear Power Plant of Pacific Gas and Electric Company, (PG&E). Initial conceptual efforts had begun earlier with Southern California Edison Company (4).

The expert system prototype uses PLEXSYS to integrate key elements of the tagout planning problem including:

1. Representation of the components and their behavior;
2. Relations between the states of individual components, subsystems and systems,
3. Representation of Tech Spec constraints on system functions, and
4. Timing of planned maintenance events.

The prototype system has been implemented on Texas Instruments Explorer and MicroExplorer systems. However, PLEXSYS is also supported at present on Sun, Symbolics, and IBM RT Workstations, and a version for personal computers based on the Intel 80386 microprocessor is currently under development.

3. Software Environment and Approach

3.1. The PLEXSYS Tool for Building Power Plant Expert Systems

The PLEXSYS concept is motivated by the idea that the description and understanding of power plant systems centers on graphical forms such as piping and instrumentation diagrams (P&IDs) and electrical line diagrams. Such diagrams define a graphics-based "model" of plant knowledge that is common to many applications, including the analysis of system reliability, the evaluation of valve and component configurations during operation and maintenance, and the predictive analysis of operational transients and accidents. The model serves as a central core of plant knowledge that can be used repeatedly as the basis for expert systems directed toward various application areas.

PLEXSYS provides a software framework within which power plant systems knowledge can be characterized and used directly in terms of schematic diagrams. PLEXSYS provides a model editor that allows the user to manually construct and modify graphical models of hydraulic, electrical, and mixed systems. Alternatively, with a planned software interface, full page P&IDs already existing on a Computer Aided Design (CAD) system could be ported to PLEXSYS and used as the basis for a plant model.

3.2. Conceptual Design of PLEXSYS

The PLEXSYS Software Development System provides an engineering tool for rapidly representing and analyzing plant systems. The PLEXSYS working environment emphasizes the direct use of schematic diagrams for designing and analyzing hydraulic, electrical and instrumentation diagrams. The PLEXSYS Development System is different from contemporary Computer Aided Design (CAD) systems in that more knowledge of the plant environment is included directly in the schematic drawing. This domain knowledge is used to assist plant personnel in designing and working with schematic drawings.

The basic components of the PLEXSYS system are described in terms familiar to plant personnel: valves, tanks, motors, pipes and pumps among other components. These elementary components are more than just simple pictures on a schematic - they have the ability to encapsulate all of the knowledge that describes the constituents of an actual component and more importantly, how it behaves as a part of a functioning system. A major design principle of the PLEXSYS system is that components can be combined into systems using this information. These systems can themselves then be manipulated as single units that can be combined with other units, components or systems to build up higher level systems at any number of levels. In principle, an entire plant can be represented in this fashion, with elementary components composing the lowest level.

Both the Plant Model Editor, the core of the PLEXSYS development package, and separate analysis packages facilitate representation of the hierarchical nature of the plant design. For the Model Editor this means a user can look ever deeper into the design from the top, while for the analysis packages, this means that during information processing, subsystems are opened and inspected as necessary.

Users are given the ability to specify their own elementary components and include them in user component libraries. These supplement the standard components provided by the PLEXSYS default environment. A user's library of components would automatically inherit the standard PLEXSYS underlying functionality. More or different functionality may be defined by the user. The user's component library may also contain specialized knowledge for connecting components, in addition to the standard component connections in PLEXSYS.

3.3. Full User Access to KEE

PLEXSYS, the specialized process plant toolkit, is implemented in the more general software environment called Knowledge Engineering Environment (KEE). KEE is a powerful software environment for building and delivering expert systems and is available on many hardware platforms. PLEXSYS architecture allows the users to use the full power of KEE and LISP. The features that are most widely used by PLEXSYS and are available to users are:

1. The KEE knowledge bases and inheritance structures,
2. The KEE representation, reasoning, and interface systems,
3. The PLEXSYS knowledge bases of graphics, standard libraries of components, and available connections,
4. The PLEXSYS plant model editor and analysis packages, and
5. The PLEXSYS' user defined component libraries and models.

3.4. KEE Resources for Developing a PLEXSYS Application

Application designers should make full use of the KEE resources when imparting new underlying functionality to the components or implementing new analysis methods. Dynamic behavior can be imparted to the plant models by using either rules or object-oriented software which incorporates the functionality of KEE to manipulate Knowledge Bases (KBs), Units, and Slot values. The major capabilities of KEE are summarized below:

1. A frame-based knowledge representation that is fully supported by rules and LISP procedures. The emphasis on frames facilitates representation of a complex domain by allowing it to be decomposed as a hierarchy of objects at varying levels of detail (abstraction). With each object is associated a number of *Slots* that characterize the objects' concrete attributes, its distinctive behavior, and procedures which it may interact with other objects.
2. A modularized rule system *(KEE Rulesystem3)* with forward and backward chaining and an assumption-based truth maintenance system that evaluates the knowledge base for internal consistency.
3. Graphical representation that can be dynamically updated based on current values of important object attributes. Graphics tools include *ActiveImages* which can be used to develop user interfaces, *KEEPictures* which define and modify low-level bitmap representations, and *Common Windows* which provide the windowing facility.

4. Active slot values that monitor the values of key object attributes. When predetermined conditions or value ranges are detected, the active values may trigger alarms, initiate a procedure, or stimulate other kinds of object behavior.

5. A sophisticated reasoning system, called *KEE Worlds* that performs hypothesis testing for a wide range of contexts including heuristic search and other applications.

6. Interfaces with other programming languages such as LISP and C and communication capabilities for linkage to several standard databases.

PLEXSYS is based on KEE (3), IntelliCorp's Knowledge Engineering Environment, and the full range of KEE functionality is available to support PLEXSYS applications. For each graphical model, PLEXSYS builds a KEE knowledge base that describes all of the component objects in terms of their individual attributes and mutual interconnections. KEE itself can then be used to build into the knowledge base additional object relationships, object behavior, and rules.

PLEXSYS also includes a Network Inspector that analyzes the model to determine available flow paths, valve closures required for isolation and maintenance of components, and other information needed to support applications. Finally, general features of KEE facilitate construction of a customized interface to serve the end user.

4. Review of Model-Based Reasoning Approach

PLEXSYS has been based upon the more general model-based reasoning paradigm, under which the problem solving knowledge base and the model knowledge base are separate, each containing its own specific type of knowledge. This paradigm's characteristics are that:

- Models are specified in terms of structured objects, object behaviors, and their relationships to other objects, and

- Problem solving procedures make reference to previously-developed domain models as the basis for performing specific kinds of analyses.

This paradigm has several benefits:

- A common model is available for use by all analysis applications.

- Development of the domain knowledge base proceeds more quickly.

- Configuration management is greatly simplified, as updating and maintaining information need be done only in the domain model.

- Multiple views of the same knowledge base are possible. For example, a pump can be viewed simultaneously as an hydraulic object in the context of a P&ID, and as an electric motor with the context of the complimentary electrical diagrams.

This approach is most effectively employed if the model includes not only the graphics model produced directly by the PLEXSYS model editor, but also any additional structure or rules that will apply across several applications.

5. Model Development

The prototype model consists of three parts:

1. The basic component layout taken directly from the P&ID,

2. The definitions of important systems and functions, and their relationships with the individual components, and

3. Definition of the "administrative state" of the plant in the context of the Technical Specification Limiting Conditions for Operation (LCOs).

5.1. Basic Component Model

The PLEXSYS model editor was used to enter the P&ID for the RHR system. The model included RHR components as well as cross-references to other system P&IDs. The diagram could then be displayed as in figure 1. Plans for the future include a general interface from IGES (Initial Graphics Exchange Specification) computer aided design (CAD) files, so that many existing diagrams can quickly be installed in a PLEXSYS model.

An important point is that PLEXSYS and KEE represent each component pictured in Figure 1 as a knowledge-base object, in the true sense of object-oriented programming, that may be given appropriate attributes and dynamic behavior. Using features of KEE, each component was assigned the attributes of availability and state. The availability of each object could assume any of the values available, unavailable or unknown. However, the possible operational states depends upon the type of component. For example, a valve can be either open or closed, and a pump state can assume the values of running or not-running.

Each component in PLEXSYS is connected to the next component on the *Canvas* via ports. Each port has the attributes of *Connection-Type* and *Directionality*. These attributes are used to define the relationships between connected components and their relationships to the subsystems and systems of the plant.

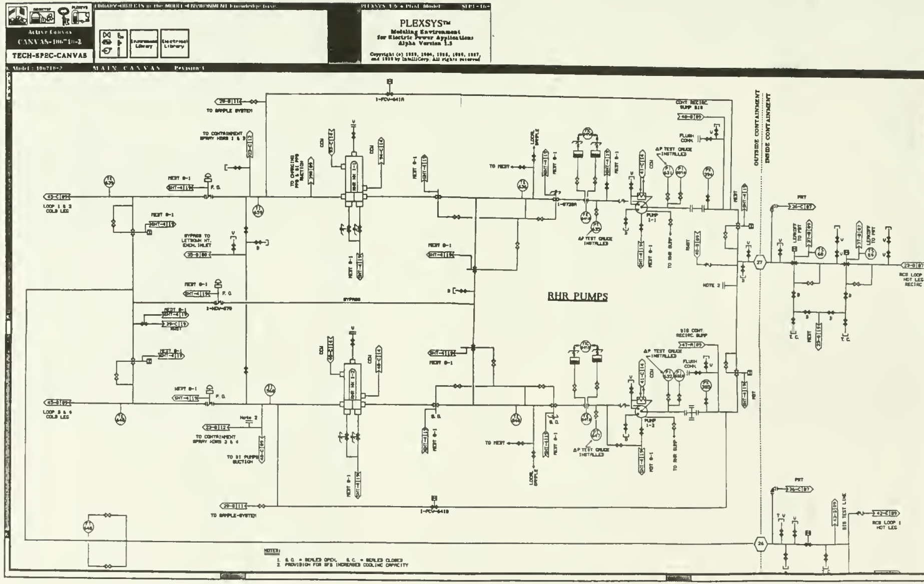


Figure 1: Residual Heat Removal System (RHR) as Modeled in PLEXSYS

5.2. Functional systems and subsystems

Additional objects are defined for the functional subsystems and systems, up to the level of the entire RHR system. The systems are assigned their own attributes of availability and operational state. The RHR system is also assigned additional attributes, such as numbers of available or operable pump trains, that relate closely to the functional requirements of the Tech Specs.

Once the basic model objects have been defined, the interdependencies between components, support equipment (such as instrumentation and power supplies) and subsystems are established, using information already available in existing plant documentation such as system fault trees.

Next, Functional Equipment Groups (FEGs) which represent the pumping trains, suction and discharge paths were defined with attributes of *Availability, State, Parts* and *Part-Of*. The first two attributes are similar to the ones that were described previously. Each FEG contains several components to perform its intended operation. As an example, the suction path from the hot-leg of the Reactor Coolant System (RCS) contains the valves 1-8701, 1-8702 and the RCS-hot-leg-4 suction path. At the same time, the valve 1-8701 is a part of the RCS-hot-leg suction path. The first relationship is described by a *Parts* and the second by a *Part-Of* attribute.

The *Parts/Part-Of*, or sometimes called *Part/Whole*, relationships are inverse of one another and are currently implemented as a part of PLEXSYS. A user must define only one of these two relationships, and the inverse is automatically determined. These *Part/Whole* relationships between different levels of model objects are summarized in Figure 2.

Note that the structure in Figure 2 relates the highest level system functions (e.g., RHR-PUMP-TRAINS) to individual components (e.g., Valve # 1-8724B) and finally to the lowest level of common support systems (e.g., Instrument Channel III). The only limit to the depth of this structure is an arbitrary grain size that is determined by the user.

This structure thus propagates a change in the availability of a low level component to that of the entire system. As an example, for each RHR loop to be considered "AVAILABLE" requires at least one suction path, pumping train (including heat exchanger), and discharge path to be "AVAILABLE". Each subsystem also requires critical instrumentation, power sources and other support systems to be "AVAILABLE".

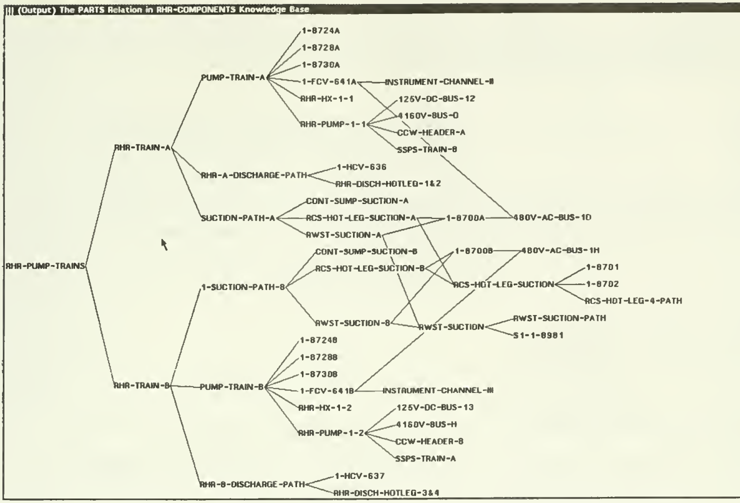


Figure 2: Part-Whole Relationship Between the RHR Subcomponents and Support Equipment

5.3. Representation of Technical Specifications

Even though the prototype model explicitly considers only the RHR system, the Tech Spec requirements for the RHR system are conditioned upon the state of other plant systems, such as the Reactor Coolant System (RCS), and upon controlled inputs such as Reactor Mode. For this limited scope prototype, such information must be supplied by the user as external boundary conditions. As the scope of a model grows to encompass a larger portion of the plant, this information is maintained internally within the model itself, and raw data may be obtained by direct access to the plant process computer and maintenance databases.

The boundary conditions for the RHR system are defined by the Tech Specs to include: Reactor Mode, Numbers of Operable RCS loops and Steam Generators, Reactor Water Level (RXWL), and Average Temperature (Tavg) for the primary loop.

These LCOs of the plant Tech Specs were implemented in the KEERuleSystem-3, in the form of "English-Like" structures called Well Formed Formulas (WFFs). WFFs are intended to be easily read and understood by an average computer literate person. An example of a WFF is:

```
(The mode of the reactor is 5).
```

WFFs are the basic elements that are used in forward and backward chaining reasoning in KEE (3).

Figure 3 presents in raw form a typical LCO, entry #3.4.1.4.1 for the Diablo Canyon RHR system. This LCO applies only if the system is in cold shutdown state (mode 5), with all RCS loops filled. The LCO requires that for time periods in excess of two hours i) one RHR loop be operating and ii) either one RHR train be operable (available) or at least two steam generators have adequate water level for heat removal. For shorter periods of time, the requirements may be relaxed.

REACTOR COOLANT SYSTEM

COLD SHUTDOWN - LOOPS FILLED

LIMITING CONDITION FOR OPERATION

3.4.1.4.1 At least one residual heat removal (RHR) train shall be OPERABLE and in operation*, and either:

a. One additional RHR train shall be OPERABLE#, or

b. The secondary side water level of at least two steam generators shall be greater than 15%.

APPLICABILITY: MODE 5 with reactor coolant loops filled##.

ACTION:

a. With one of the RHR trains inoperable and with less than the required steam generator water level, immediately initiate corrective action to return the inoperable RHR train to OPERABLE status or restore the required steam generator water level as soon as possible.

b. With no RHR train in operation, suspend all operations involving a reduction in boron concentration of the Reactor Coolant System and immediately initiate corrective action to return the required RHR train to operation.

Figure 3: Typical Tech Spec LCO for the Diablo Canyon RHR System

All the applicable LCOs for the RHR system are characterized succinctly in Figure 4. This figure provides the basis for constructing rules that describe the Technical Specifications. Note that lines 12 through 18 of figure 4 summarize the 7 subcases of the LCO described above. In this figure, each row is numbered according to the actual Tech Spec, and each column represents the parameters that govern whether that LCO is *"Fired"* or not. *Firing* an LCO means rejecting the requested MWR because a licensing requirement would be violated. The set of KEE rules corresponding to LCO # 3.4.1.4.1 is shown in Figure 5.

6. Maintenance Tagout Planning Application

It should be emphasized that the model described in Section 3 can be defined independent of the particular application. The utility of the basic model thus extends beyond the context of the tagout planning application and may be used in other applications such as diagnosis and/or alarm monitoring.

6.1. Description of application

The objective of the prototype expert system is to identify and resolve conflicts between proposed maintenance actions and requirements of Technical Specification Limiting Conditions of Operation (LCOs).

It is assumed that a queue of approved maintenance work requests (MWR) exists and that the maintenance planner wishes to augment the queue by proposing a single maintenance action that involves removing one or more components from service for some period of time, known as the "proposed time window". The expert system assists the planner with incrementally augmenting the queue of maintenance requests, while ensuring that no LCOs are violated by any tagouts implied by the proposed maintenance action. The queue itself could be included as part of the system, but it would more likely be maintained as a mainframe database to be accessed by the system.

The system considers the proposed maintenance request together with previously approved maintenance requests to determine the functional state of the plant system during the proposed time window. This functional state is then compared with all relevant requirements of the LCOs, which in turn depend upon the plant mode and other conditions planned for the proposed time window. Should all LCO requirements be satisfied, the planner is notified of compliance so that the proposed action may be added to the approved queue.

However, when conflicts are identified, the system will provide explanations that help the planner identify acceptable alternatives. Such explanations include descriptions of the relevant LCOs and specific indications of how the proposed component maintenance action would violate the LCO requirements, or if any of the LCOs were violated, what are the action items that the operators must follow.

pgetechspecs

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	LCO	MODE	RCS OPERATING	RCS AVAIL.	RHR OPERATING	RHR AVAIL.	TIME	RWST	CT SUMP	T avg (F)	WATER LEVEL	RCS STATUS	SG'S W/>15% WL
2													
3	3.4.1.3-CASE-A	HOT-SHUTDOWN	'=1	'=1	'=0	>=1							
4						'=0	<=24						
5													
6	3.4.1.3-CASE-B	HOT-SHUTDOWN	'=0	'=0	'=1	'=2							
7					'=0	'=0	<=1						
8				'=1	>=1	>=1							
9													
10	3.4.1.3-CASE-C	HOT-SHUTDOWN	>=1	>=2									
11													
12	3.4.1.4.1-CASE-A	COLD-SHUTDOWN			>=1	'=2						FILLED	<2
13					'=1	'=1	<=2						
14					'=0	>=1	<=1						
15													
16	3.4.1.4.1-CASE-B	COLD-SHUTDOWN			>=1	>=1						FILLED	>=2
17					'=0		<=1						
18													
19	3.4.1.4.1.2	COLD-SHUTDOWN			>=1	'=2						NOT-FILLED	
20					>=1	'=1	<=2						
21					'=0	>=1	<=1						
22													
23	3.5.2	AT-POWER				'=2		YES	YES				
24		STARTUP											
25		HOT-STANDBY											
26													
27	3.5.3	HOT-SHUTDOWN				'=1		YES	YES	<350			
28													
29	3.9.8.1	REFUELED			>=1	>=1					>23		
30					'=0	'=0	<=2				>23		
31													
32	3.9.8.2	REFUELED			'=1	'=2					<=23		
33					'=0	'=2	<=1				<=23		
34													
35	DEFUELED	DEFUELED											

Figure 4: Spreadsheet Summary of LCO Cases Applicable to the RHR System

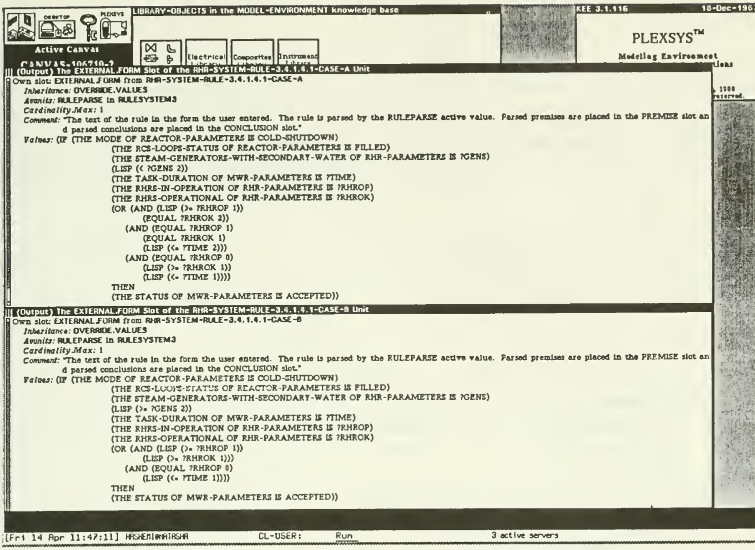


Figure 5: KEE Rules Corresponding to LCO # 3.4.1.4.1

Figure 6 summarizes the major functions of the expert system. Based upon the reactor mode and other "boundary conditions" (i.e., outside the boundaries of the current model), the Tech Specs define the minimum requirements for the RHR system. The PLEXSYS Network Inspector, through its tagout boundary analysis option described in Section 6.2, determines the additional valves that need be removed from service in addition to the maintenance work request. For the proposed component configuration, the domain model determines the actual system availability and state for comparison against the Tech Spec requirements.

Each maintenance work request identifies the component, the general class of activity, and a time window characterized by a starting and stopping time. In a full-scale application, this system would be used for planning time periods in the future. However, for the present prototype demonstration, each time window is assumed to begin at the present time, so that it is fully characterized by a single time value that defines the duration of the activity.

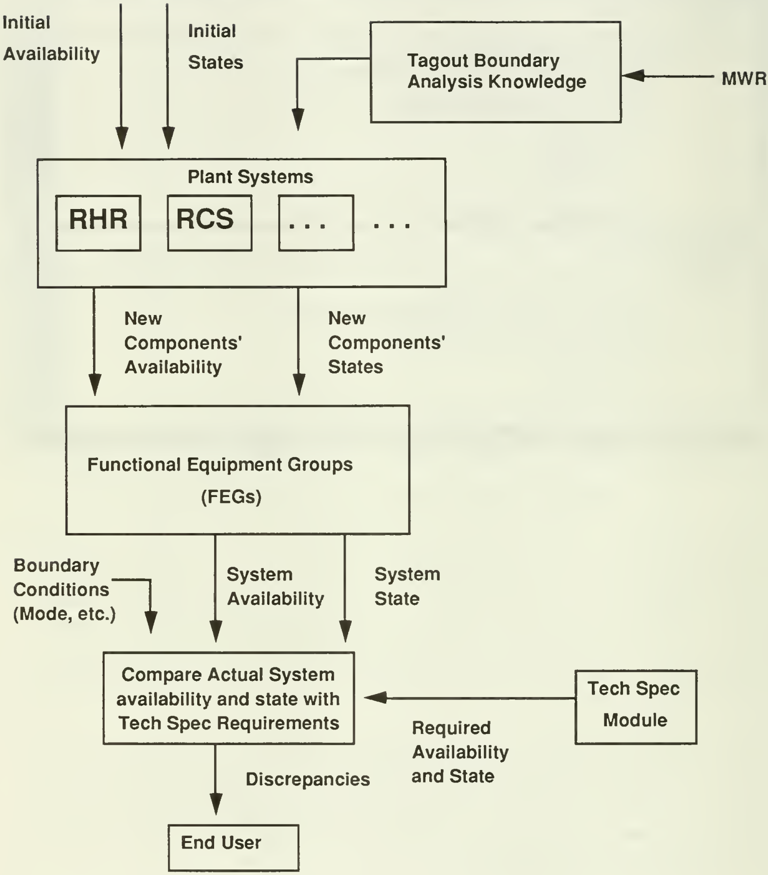


Figure 6: Elements of the Expert System Analysis for Tagout Application

6.2. User Interface

Prior to designing a user interface, the developer must first clearly determine 1) any processes to be controlled and the types of inputs to be supplied by the user and 2) the output information that is to be displayed to the user. The KEE ActiveImages features provide predefined functions that can be used to supply input values and commands via mouse and menu operations and to present output information in a variety of forms such as text, meters, and bar graphs.

The ActiveImages features of KEE have been used to construct a customized user interface, shown in Figure 7, for the tagout planning application. The interface consists of several windows for controlling the expert system and observing its output. Each entry in these windows can be accessed by pointing with the mouse.

The Plant Conditions window is used to review or modify the major plant boundary conditions, such as the operating mode or the number of active coolant loops. These boundary conditions can be changed to evaluate plans for changing the operating state of the plant in terms of their effect on Tech Spec constraints.

The user wishing to evaluate a proposed MWR mouses on the appropriate control panel item; the system then prompts the user to identify the component to be isolated and the type of isolation (e.g., hydraulic or electrical). The PLEXSYS Network Inspector searches the network of pipes and instruments to identify the isolation boundary and all affected components, and the boundary is highlighted for the user's inspection. Following the user's confirmation, the system marks all the affected components as "UNAVAILABLE" and updates the availability of the subsystems and the overall RHR system.

Next, the user selects "Run Tech Specs" to retrieve and activate the Tech Spec rules. If the request is rejected, as in Figure 7, more detail about violated LCOs will be supplied in the user dialogue window, by mouse clicking on the rejected LCO. This functionality is added to serve as a guide to the user in submitting a modified or alternative MWR.

6.3. Tagout System Operation -- Examples

This section provides a simple sequence of examples illustrating the types of requests and information available from the prototype system.

Consider a starting point (Fig 7) in the cold shutdown mode 5, with both RHR loops operational, but

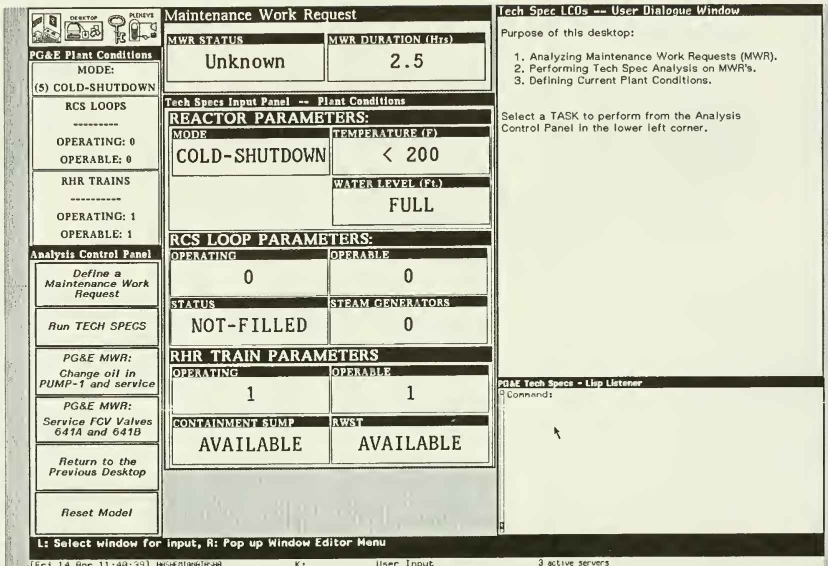


Figure 7: Maintenance Work Request Control Panel for the RHR System

with all steam generators empty. A proposed maintenance work request would require the main RHR pump to be isolated for two and a half hours for an oil change. Since the entire loop would be down because of this activity, the maintenance staff could consider adding the valve 1-8728A to the components being inspected or maintained during that time, since that valve will not extend the isolation boundary to the second loop.

Figure 8 shows the system response following submittal of this MWR. Because one of the RHR pumps would be deenergized for more than two hours, LCO #3.4.1.4.1 and #3.4.1.4.2 have been violated, and the MWR is thus rejected. Assuming that the maintenance action could be speeded, an alternative MWR could be proposed for the shorter time duration of two hours. As shown in Figure 9, this alternative plan satisfies all the LCOs, and the Tech Spec evaluation produces an acceptable result.

7. Summary and Conclusions

This paper illustrates how features of PLEXSYS and KEE can be used to build an application-specific expert system for a power plant application. This example also emphasizes the division of expert knowledge between the permanent model, which can be reused for many applications, and the knowledge that is specific to the immediate application.

The greatest benefit of PLEXSYS-based modeling and analysis is that all changes, either to the physical or "administrative" (i.e., Tech Specs) model can be reflected in the knowledge base with a minimum effort. By performing such updates on the central model, the rest of the system becomes aware of the changes automatically, and the issue of configuration management control is greatly simplified. The model can be extended as needed to include more plant systems in a more extensive application. Furthermore, the model is directly usable for a variety of other applications, including reliability analysis, plant design modifications, malfunction diagnosis, and analysis of alternative scenarios for planning and scheduling.

The prototype system described in this paper can easily be linked, using a terminal window and either a modem or an Ethernet network, to mainframe-based data bases and other application software such as planning and scheduling algorithms. Results of the PLEXSYS analysis can easily be formatted for compatibility with the mainframe programs and then uploaded to provide input for plant-wide analysis.

The prototype system can be integrated with the scheduling system to create plans for maintenance activities during the plant refueling outages and unanticipated shutdowns. Such an integrated capability could be extremely powerful in quickly adjusting to contingencies or unanticipated problems, such as unavailability of essential spare parts or equipment failures. The schedule could be revised very quickly with the potential for reducing overall down time during a forced outage and under the changing

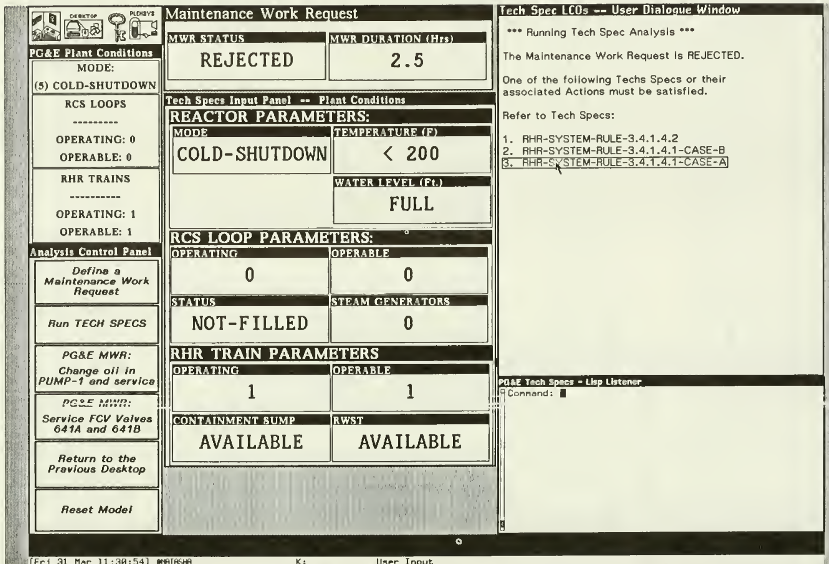


Figure 8: Failure of a Maintenance Work Request due to Violation of LCOs

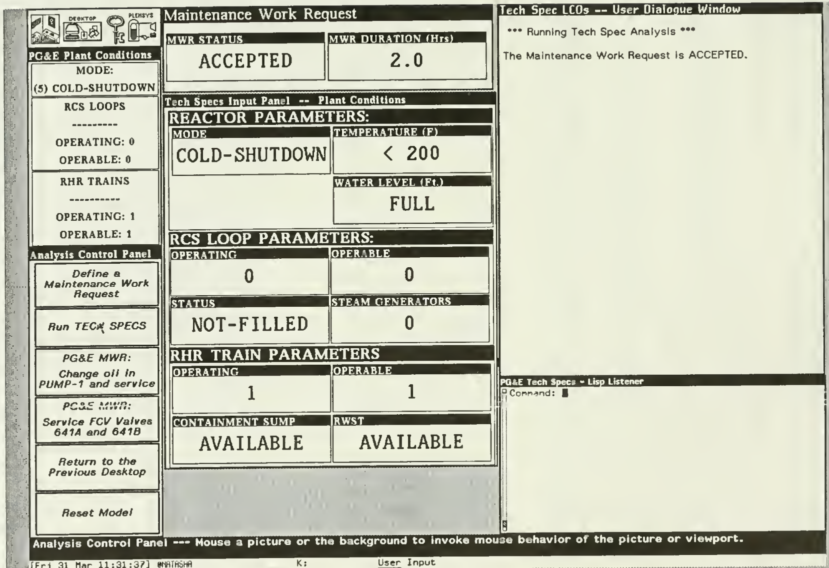


Figure 9: Success of the Maintenance Work Request by Compliance with all LCOs

constraints faced during a planned outage.

8. References

1. S. Hashemi, PLant EXpert SYStem (PLEXSYS), A Model-Based Reasoning Approach to Plant Modeling, Proc. American Institute of CHemical Engineers (AICHE), Houston, Texas, April 1989.

2. S. Hashemi et al., PLEXSYS PLant EXpert SYStem Development Environment Description and User's Manual Version 2.0, EPRI Report in preparation for Project No. RP2582-03, March 1989.

3. KEE Users Guide, KEE Version 3.1, IntelliCorp, Inc., Publication No. K3.1-UG-1, Mountain View, CA, May 1988.

4. J. H. Munchausen and K. III. Glazer, An Expert System Technology for Work Authorization Information Systems, EPRI Seminar: Expert Systems in Power Plants, Boston, May 1987.

5. G. III. Finn et al., An Expert System for Technical Specifications, EPRI Seminar: Expert Systems in Power Plants, Boston, May 1987.

6. H. T. Roman et al., Expert Systems Applications at PSE&G - A Task Force Approach, EPRI Seminar: Expert Systems in Power Plants, Boston, May 1987.

Model-based Reasoning Technology for the Power Industry

R. A. TOUCHTON and N. S. SUBRAMANYAN
Technology Applications, Inc.
6621 Southpoint Drive North, Suite 310
Jacksonville, Florida 32216, USA

J. A. NASER
Electric Power Research Institute
3412 Hillview Avenue
Palo Alto, California 94304, USA

ABSTRACT

Model-based reasoning refers to an expert system implementation methodology that uses a model of the system which is being reasoned about. Model-based representation and reasoning techniques offer many advantages and are highly suitable for domains where the individual components, their interconnection, and their behavior is well-known. Technology Applications, Inc. (TAI), under contract to the Electric Power Research Institute (EPRI), investigated the use of model-based reasoning in the power industry. During this project, a model-based monitoring and diagnostic tool, called ProSys, was developed. Also, an alarm prioritization system was developed as a demonstration prototype.

INTRODUCTION AND TERMINOLOGY

As a part of NASA's Systems Autonomy Program, personnel at Kennedy Space Center (KSC) have developed a prototype for performing real-time, knowledge-based system monitoring, system diagnosis, control, and reconfiguration. This system is called Knowledge-based Autonomous Test Engineer (KATE). Many of the technical barriers addressed and overcome by the KSC effort are currently R&D issues within the electric power industry. Research Project RP2902-1, Nuclear Power Applications of NASA Control and Diagnostics Technology, analyzed the NASA technology and identified techniques useful in the electric power industry. Model-based reasoning techniques were refined and reimplemented in ProSys. An application was selected after plant interviews and a demonstration prototype was built to illustrate the benefits of this technology.

This paper describes ProSys, the techniques used in ProSys, and the general course taken by the project. First, we define certain words and phrases that are used in this paper. The next section describes model-based reasoning and object-oriented programming techniques that were used in the project. Then, the progress of the project is described in detail including the objectives, the main elements, the development of ProSys, and the development of a demonstration prototype. This is followed by the conclusion.

We define below certain terms that are used in the rest of the paper.

System or **Computer System** refers to ProSys, applications built using ProSys, or in general, other computer software systems that are used for monitoring, diagnostics, and/or control.

Real System or **Physical System** refers to the real-world system that is being monitored and in which problems are being diagnosed.

Model is the representation of the real system inside ProSys.

Simulation is a copy of the model used instead of the real system to supply measured values for the ProSys diagnoser. ProSys needs measurements from the real system to perform diagnosis. Since it is not possible to "hook up" to a real system during development and testing, the simulation provides the needed measurements. Faults can be created in the simulation by the user and subsequently diagnosed by ProSys. There is no link between the simulation and the diagnoser and hence the diagnoser has no access to the failure information.

Sensors are the real-world measuring devices and their representations in the model.

Discrepancies are the disagreement between the values coming from the sensors in the real system (or the simulation) and the expected values of sensors in the model. While monitoring the real system, ProSys uses the discrepancies to recognize that there is a problem with the real system.

MODEL-BASED REASONING AND OBJECT-ORIENTED PROGRAMMING

Model-based Reasoning

Expert systems have evolved from simple rule-based systems to object-oriented frame-based systems. Simple rule-based expert systems provide only limited capability to model and explore problems. While the human expert may use structural and functional domain knowledge for solving a problem in a rule-based system, such knowledge is often entangled with problem-solving heuristics. Such knowledge is termed "compiled" or "implicit" knowledge and is of limited use. On the other hand, the frame-based environment provides a framework for building "free-standing" models of problem areas which can be analyzed and used in a variety of ways. Such a model is easier to maintain and extend and thus has a larger life-span than that provided by totally rule-based systems. Further, in cases where the processing and use of the model can be generalized, the system will be able to solve problems not explicitly thought of before.

Modeling is the process of building computational equivalents of the objects in the problem domain. Models that are rich enough to be useful as problem-solving tools can then be analyzed using various techniques appropriate to different applications. Some advantages of model-based expert systems are as follows:

- Adaptability - As mentioned before, the model that is built is "free-standing." This refers to the explicit nature of the knowledge contained in the model. The knowledge does not depend on any particular application, only on the physical system itself. Such adaptability increases with the integrity of the model (i.e., how closely it defines the system). In other words, this problem-solving approach affords different perspectives to solve different problems with the same knowledge base.

- Increased Life Cycle - The model itself can be readily modified and extended to reflect changes and growth in the problem domain. Thus, the system may be fine-tuned by incrementally refining and enhancing the model.

- Reduced System Cost - A single model with multiple interpretations and uses leverages the development and maintenance costs. The ease of adaptability and the increased life cycle are manifested as reduced life cycle costs. Since many applications of this technology are anticipated, this advantage is especially important.

- Verifiability - Explicit models are easier to verify because they represent fundamental knowledge about the system.

- Potential for Handling Unexpected Situations - Since the knowledge is "uncompiled" and free to be interpreted, there is greater potential for handling of situations unanticipated by the expert system developer/modeler.

- Portability - Frame-based environments are available for most AI and conventional hardware. This advantage will permit systems based on ProSys technology to be ported to different hardware with minimal work. (A further advantage of the ProSys technology is that it was developed using Common LISP which facilitates porting to various computer systems. Thus, applications may be moved to the computer hardware which best accommodates budget limitations, speed requirements, and size of the application.)

Object-oriented Programming

Object-oriented programming is an evolution of programming. Much like the structured programming concepts introduced by languages like Pascal, object-oriented programming tools offer facilities that make some programming tasks easier and more natural. In object-oriented programming, each concept or entity in a problem is represented by a "software object" inside the system. This software object stores all data associated with that entity and procedures that can be performed on or by that entity. Thus, the software object contains the entire definition of the entity and so contributes to the modularity and expressiveness of the system. Also, such software objects can be linked together and can inherit data and procedures from one another. This reduces the redundancy in the storage of similar data and procedures because they can be stored once and then inherited whenever they are needed.

The object-oriented programming paradigm is very appropriate for model-based reasoning. Building explicit models involves defining an object for each component. Also, since many components are similar, it is useful to define the component once and then inherit the properties in actual component "instances." In this project, an expert system environment called KEYSTONE was used to provide the object-oriented facilities in the form of a frame language.

PROJECT PHASES AND RESULTS

Project Objectives

The overall objective of this project has been to explore the applicability of this NASA technology to problems encountered in the electric power industry. The original work objectives can be further divided into the following:

- to dissect and assess the KSC technology
- to identify and prioritize utility application possibilities
- to develop a demonstration prototype of an application which will help to communicate the technology and its problem-solving capabilities to utility industry personnel

Project Elements

This project consisted of several distinct, but interdependent, elements as depicted in Figure 1. This subsection defines each element and summarizes the results of the project for it.

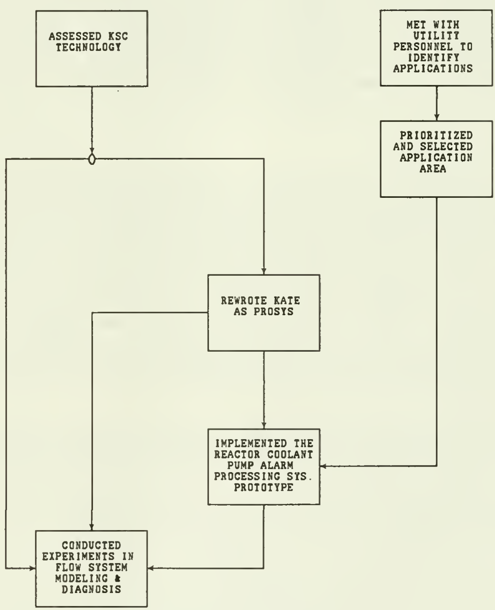


Figure 1. Project Elements

The first element of the effort was for the project team to learn and evaluate the NASA technology in order to identify its applicability and use in solving electric power industry problems. Thus, the TAI Project Team spent a considerable effort assessing the KATE software and methodology. This effort also included extensive discussions with the KSC Team that developed, and continues to enhance, the NASA prototype.

Another important element of this effort was to gain utility input regarding the areas where integration of the NASA technology might prove beneficial in the nuclear power industry. Thus, ten utilities were visited and given a project briefing followed by a brainstorming session. Forty-four potential applications were identified and organized into four categories: on-line control and monitoring systems, on-line advisory systems, off-line advisory systems, and "other." Based on the utility discussions, each application was assigned ratings in terms of attributes such as level of support, priority level, and other considerations.

In conjunction with the utility dialogues and KATE assessment, each of these areas (as well as any new ones suggested) were explored to quantify the enhancement of electric power industry capability, functionality, and/or performance. An assessment was made as to how well the NASA Systems Autonomy core technology could fill needs of the utilities. The applications were prioritized based on their estimated cost/benefit, risk, and utility support. The four application areas receiving the highest evaluation ranking were:

- Alarm Screening/Intelligent Annunciators
- On-line Thermal Performance Advisor
- On-line Technical Specifications Monitor/Advisor
- On-line Root Cause Analyzer

The first of these was selected as the subject of the demonstration prototype. In its current state of maturity, KATE can only deal with a limited subset of utility needs.

The project also included a software development effort which was conducted on three planes. First, there was identified a need to make the NASA software more generic and more tuned to ultimate users in the electric power industry. Therefore, KATE was rewritten as ProSys, a user-friendly "shell" for creating and using KATE-style models. Next, an alarm processing demonstration prototype was developed based on a simplified reactor coolant pump seal water injection system. Finally, an experiment was conducted to explore alternative diagnostic techniques which would not be subject to so many of the limitations incurred using the original KATE method. A qualitative reasoning technique was shown to offer considerable promise for multi-path flow systems.

PROSYS - THE TOOL

ProSys System Description

ProSys is a model-based diagnostic system that is built on basic principles of troubleshooting, such as cause and effect, and not on heuristics derived from experience. Models in ProSys store knowledge about the structure and function of the system being diagnosed. ProSys uses this knowledge to draw inferences about the current state of the system. By comparing the values reported from the field and the expected state of the system, ProSys is able to hypothesize and confirm failures in the components of the system.

ProSys falls under a class of computer systems called knowledge-based or expert systems. Knowledge-based systems are different from conventional software systems in that they have some features which facilitate the creation of more adaptive and extendable programs. One of the features is the separation of the declarative (factual) portion of the program from the procedural portion. Since the solution procedure does not change too much between different applications, it is possible to develop different applications just by changing the declarative portion.

For example, a diagnostic procedure may be divided into the major rules of diagnosis, and then declarative knowledge about the physical system being diagnosed. To diagnose a different physical system, provided the rules are general enough, the user need only replace the declarative knowledge about the physical system. Such explicit, declarative knowledge is called the "model."

ProSys Architecture

The architecture of ProSys is shown in Figure 2. ProSys is built using KEYSTONE, which is an expert system development environment that provides a frame language and other facilities for object-oriented programming. Using these facilities, each component in a model can be represented by one object inside the system. Such "software" objects can be connected together to form an entire system model. ProSys stores the models and other system information in collections or groups of software objects called knowledge bases (KBs). Thus, the ProSys KB in the figure stores knowledge that is common among the models. It also incorporates a diagnostic algorithm which diagnoses faults in the model based on sensor information reported from the real system.

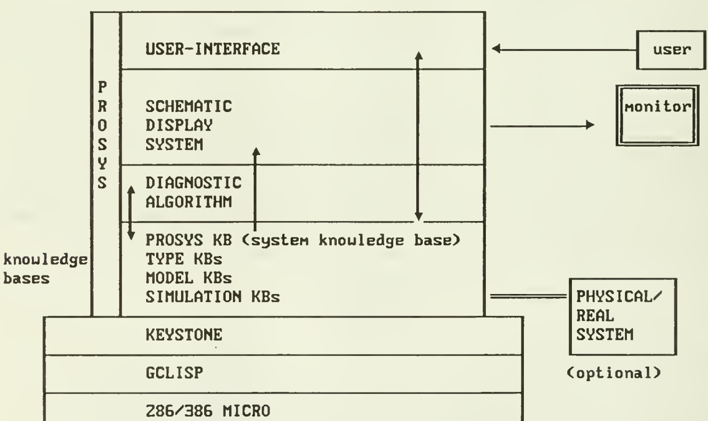


Figure 2. ProSys Architecture

The schematic display system displays a schematic diagram to a window on the screen. This diagram is used to provide a visual display of the model and the values at the outputs of each component in the model. It is also used to connect and disconnect components during model-building. The ProSys interface is very user-friendly and uses menus and prompts to guide the user through model-building and diagnosis activities.

KEYSTONE is written using Golden Common LISP which runs on the widely available 80286 and 80386-based microcomputers. Golden Common LISP is an implementation of Common LISP and the source code is quite portable across different machines.

Model-building in ProSys

A model of the physical system is created using the ProSys software. This model supplies the necessary knowledge to ProSys so that it may reason about the physical system and its behavior. Since ProSys is an experimental system for which portability and low cost are very important, it does not yet interface with any physical system. Instead, a copy of the model (SIMULATION) is used to simulate failures and ProSys tries to diagnose those failures based on the simulated measurement values generated by the SIMULATION. It is expected that ProSys's powerful monitoring and diagnostic capabilities will also be brought to bear on plant simulators and actual plant equipment.

In order to formalize model-building activity in ProSys, certain constructs have been identified. They are components, commands, measurements, and alarms. Components are the functional parts of the system such as valves, pumps, control circuitry, etc. Commands are user inputs to the physical system (like the position of a manual valve). Measurements are the sensor outputs of the system. Alarms are representations of the individual alarms in the system's alarm panel and contain the associated measurement setpoints or logic (e.g., HIGH-REACTOR-TEMP (alarm) is TRUE when RCS-TEMP-1 greater than 900F).

Every object in the model is based on one of these constructs. A ProSys model is built by creating the components, commands, measurements, and alarms and by establishing connections between them. ProSys model-building facilities are described in detail in [9], Volume III.

Diagnosis in ProSys

The strategy behind ProSys is to compare the behavior of a real-world system (or the SIMULATION) to that of a software model that is designed to closely represent the real-world system. For this, ProSys must have a knowledge of what control inputs were fed into the real system. These control inputs are called "COMMANDS." Also, for monitoring, the real system measurements should be reported from the sensors.

ProSys detects a problem when there is a discrepancy between the field measurements and the measurements predicted by the software model. It then explores its software model (just as an engineer would) to determine which component failure would account for or cause the set of field measurements. This process is one of systematic analysis using the structure of the model and the function of the various components. First, the list of components is pruned to remove those components which cannot influence the discrepancy. Then the failure of each of the remaining components is hypothesized. The failed value (for hypothesis) is obtained by back-calculation from one of the field measurements. The measurements are compared once again, with the "hypothesized failure" in place, to see if they are consistent. If the measurements in both the real system and the model are the same, then the "hypothesized failure" is a possibility, else the component is removed from consideration. See [8] for a complete description of the ProSys diagnoser.

Thus, ProSys reacts to discrepancies between the software model and the real world, and finds the cause for the discrepancy by systematically reasoning upon the model until the variance is accounted for. This approach is well-suited for identifying malfunctions in physical systems.

ProSys User Interface

ProSys makes extensive use of menus and icons to provide a friendly user-interface. Icons are small pictures on the screen which represent a system object or function. They are usually mouse-sensitive; that is, by placing the mouse cursor on the icon and clicking the mouse button(s), the user can accomplish some related functions. Typical functions might be as simple as displaying a description of the object described by that icon or as complex as invoking a function that changes the position of the object on the schematic or its value.

ProSys has a diagnoser-trace window which is scrollable up and down. The diagnoser sends text strings to this window as it goes through the diagnostic process. The contents of this window are available for perusal until the diagnoser is invoked again. The trace can also be written to a disk file and then sent to the printer for a hardcopy.

The schematic display facilities of ProSys allow the user to display any model in a schematic form, similar to a P&ID (Piping and Instrumentation Diagram). The schematic display system is built to use the icon definitions and the connection information stored in the model. Also, ProSys can plan a layout on its own through a process referred to as recalculating the schematic. Since this process can be time-consuming and aesthetically imperfect, ProSys offers another option for planning the diagram layout. This option allows the user to place each component on the screen by pointing to the specific position using the mouse and clicking the left button. The layout information is just a screen coordinate stored with each ProSys construct. Once a layout has been calculated or specified for a particular model, ProSys will use that layout unless the user asks to recalculate again. When there are additions to the model, the schematic system prompts the user to place the added construct at a preferred position in the schematic using the mouse.

THE DEMONSTRATION PROTOTYPE

The complexity of modern power plants and the sophistication of the computer-based systems that control them enables the monitoring of thousands of alarm points. These alarm points are typically monitored independently of one another, making it likely that a single fault will directly generate a single alarm, and indirectly generate numerous others. Such cascading alarms can quickly overwhelm the plant operations staff. The goal of an alarm processing system is to aid the operator during plant transients and off-normal events. By minimizing the amount of visual clutter that confronts the operator during transients, the alarm filtering system will improve plant performance and enhance plant safety. The alarm processing demonstration prototype developed for this project is described briefly in the following paragraphs.

Prototype System Selection and Modeling

ProSys does not have built-in abstraction capabilities (the ability to work with coarse overview of systems) to allow modeling of systems with many components. The requirements of the alarm processing prototype application suggested finding a system that also had enough associated alarms with which to work. After examining Alarm Response Procedures from a Pressurized Water Reactor, the seal injection system in a Reactor Coolant Pump (RCP) was selected as the candidate system. The function of the seal injection system was to provide controlled inleakage into the RCP so that there is essentially zero reactor coolant leakage into the containment via the shaft.

The ProSys model of the seal injection system was limited to the major system components (e.g., seals, flow sensors). Components such as pipe segments and fittings, check valves, etc., were ignored and their resistance to flow was lumped with nearby prototype components. The main emphasis was on alarms associated with this system. The alarms deal almost exclusively with abnormal pressures and flows through system components. Most of the alarms generated in the prototype have real-world equivalents that are annunciated in the plant control room.

First, prototype objects were defined for the pump, the seals, and the pressure and flow sensors. Then, instances were created to represent each occurrence of the above-mentioned prototypes and then connected to complete the model. Details of the prototype object definitions can be found in [9], Volume II.

Alarms and their processing

Early in the project, three methodologies of screening alarms were identified. The batch mode of alarm processing would use an off-line procedure to build an alarm dependency network consisting of all the accompanying alarms that are generated by a single component failure. Then, alarms would be filtered by matching the predetermined network of alarms with the actual alarms that occur in the system. The model-based approach creates a list of possible faulty components using the system model and diagnosis. By simulating the effects of each fault, it would be possible to decide which alarm to emphasize. The final method is the use of functional relationships that can be identified from common engineering practice and from insights obtained through knowledge engineering with senior plant operations staff.

The functional relationship method mentioned above was used to assemble the network of alarms used in the demonstration prototype. Alarms were modeled as having one output, the value of which determines whether the alarm is active or not. The alarm value, in turn, is a function of some number of inputs, so in effect, an alarm resembles a measurement object with multiple inputs and a behavior which describes the activation criteria. Also, the names of secondary alarms are stored in the alarm object for specifying the functional relationship (i.e., which alarms are secondary to which other alarms). If a particular alarm is active, then all its associated secondary alarms are de-emphasized. Alarms from both the model and the simulation are displayed, and the functional relationships are used to de-emphasize the secondary alarms only in the model. Thus, the user can see, on the same screen, a set of unprioritized alarms from the simulation and another set of prioritized alarms from the model.

Work on the alarm processing application proved that it was indeed possible to model and simulate physical systems and alarms associated with these systems. It also established that functional (precursor) relationships could be represented in the model and used to prioritize alarms. This effort also raised various development and research questions with respect to the KATE technology which were examined and documented in [9].

FLOW SYSTEM EXPERIMENT

In its current state, the ProSys technology does not work well with fluid or hydraulic systems. In such systems, changes in user controls and changes in the state or health of a component have system-wide effects, and this is mainly due to the "bidirectional" nature of the components involved. The behavior of each component cannot be described just by describing its outputs as a function of its inputs; one also has to account for the fact that the input values themselves are dependent on the flow capacities of components connected to the output. Flow capacities, which represent the resistance to flow offered by a flow component, are present in all flow systems. This behavioral complexity was reduced by "teaching" ProSys about the system-wide influence of flow capacities of components. The modeling abilities of ProSys were extended to model flow capacities in each flow component and also to combine these flow capacities to calculate effective capacities at various points in the system.

The diagnoser was changed to use some fundamental flow system characteristics to qualitatively analyze the model using pressure and flow trends. This is different from the KATE/ProSys diagnoser which quantitatively generated hypothesis and simulated them to confirm their validity. The pressure and flow trends mentioned above are the differences between the values generated from the model (expected values) and the values reported from the real system (measured values). For example, if the measured value is higher than the expected value, then the trend is "increasing." The actual development of the diagnostic algorithm from basic principles is described in [9], Volumes I and II.

The flow system experiment proved the concept of quantitative simulation and qualitative diagnosis. Additional work needs to be done for applying this technique to general flow system topologies. Used selectively, this technique promises to alleviate the computational complexity of diagnosing such highly interacting systems.

CONCLUSION

In general, it was proven that given enough information about the physical system in the form of a complete model, a generic system can monitor and troubleshoot the physical system. The main advantage of such a generic system is that it is very easy to maintain and extend, because any change in the design of the physical system need only be reflected in the model.

Development of ProSys, the alarm processing application, and the exploration of new techniques to solve flow system problems was an important exercise and contributed significantly to the understanding of strengths and weaknesses of the KATE technology. Further, the effort has also produced ProSys, a user-friendly modeling and diagnosis tool that embodies all the important and proven KATE techniques to further research and development in this intriguing area of model-based simulation and diagnostic systems.

While tremendous inroads have been made in understanding the KATE technology and its limitations, further effort is necessary to apply this technology in more challenging domains. The research conducted in this phase of the project indicates that the KATE technology can be successfully applied in some selected areas. Systems with feedback and components with state need more work before KATE techniques can be beneficial and certain others, involving complex time dependencies, bidirectionality, and integral quantities, violate fundamental assumptions underlying KATE and may not ever be suitable for practical application of KATE techniques.

REFERENCES

1. Jamieson, J. R., "Knowledge-based Automatic Test Equipment," Proceedings National Conference on Artificial Intelligence, Philadelphia, August 1986.

2. Benjamin, K., "Common Sense Reasoning about Causality: Deriving Behavior from Structure," D. G. Bobrow (Ed.), Qualitative Reasoning about Physical Systems (The MIT Press, Cambridge, Mass., 1985).

3. Davis, R., "Diagnostic Reasoning Based on Structure and Behavior," D. G. Bobrow (Ed.), Qualitative Reasoning about Physical Systems (The MIT Press, Cambridge, Mass., 1985).

4. De Kleer, J., "How Circuits Work," D. G. Bobrow (Ed.), Qualitative Reasoning about Physical Systems (The MIT Press, Cambridge, Mass., 1985).

5. De Kleer, J. and B. C. Williams, "Reasoning about Multiple Faults," Proceedings National Conference on Artificial Intelligence, Philadelphia, August 1986.

6. Williams, B. C., "Doing Time: Putting Qualitative Reasoning on Firmer Ground," Proceedings National Conference on Artificial Intelligence, Philadelphia, August 1986.

7. Touchton, R. A., N. Subramanyan (Technology Applications, Inc.), and J. Naser (Electric Power Research Institute), "Expert System Diagnostics for Nuclear Power Plants," Proceedings of the 23rd Intersociety Energy Conversion Engineering Conference, Volume 1, 1988.

8. Subramanyan, N., R. C. Antinoja (Technology Applications, Inc.), and J. Naser (Electric Power Research Institute), "ProSys - Model-based Diagnosis and Signal Validation," Proceedings of the 7th Power Plant Dynamics, Control and Testing Symposium, May 15-17, 1989.

9. Touchton, R. A., N. S. Subramanyan, A. Lane, P. Hariharasubramanian, and T. D. Eldredge II, "Nuclear Power Applications of NASA Control and Diagnostics Technology," Volumes I, II, and III, submitted to the Electric Power Research Institute as part of the Final Report for RP2902-1.

Man-Machine Interface Aspects of Experts Systems in the CEGB

J. N. IBISON
Central Electricity Research Laboratories
Kelvin Avenue
Leatherhead
Surrey KT22 7SE, UK

ABSTRACT

This paper discusses various human issues related to user interfaces with reference to CEGB projects. Several projects are described in terms of the user interface issues which they highlight. This is followed by a discussion showing the way in which these issues were addressed in one particular project. The interface design process is described and the effectiveness of the techniques employed is discussed.

INTRODUCTION

The Central Electricity Generating Board is the body responsible for the generation and transmission of electricity within England and Wales. Part of the role of the Research Division within the CEGB is to keep abreast of new technology and look for improvements which can be made in terms of performance, security and safety. Expert systems are seen as a potentially valuable technology; this paper discusses some of the work done by the CEGB on the user interface aspect of expert systems.

The aim of this paper is to illustrate work on the man machine interface aspects of expert systems. The content is divided into two main sections. The first gives a fairly broad look at several systems under development and aims to give a general overview.

The subsequent section focusses on one particular project which has a significant user interface component, the R6 Interface Project. One of the particular features of this project was the importance maintaining a good working relationship with the clients, because the clients were to provide the domain expertise. This Project therefore highlights the importance of human issues. The design process for the R6 Interface is a particular theme of this paper, because it illustrates one way in which both technical and non-technical issues can be tackled together.

A DISCUSSION OF VARIOUS CEGB PROJECTS

The man-machine interface is of central importance to a wide range of IT applications, although it is perhaps only more recently that it has received the full attention due to it. The progressive realization that the ergonomic aspects of a system may completely outweigh considerations of functionality in influencing user acceptance has led to a burgeoning of interest and the emergence of techniques aimed specifically at interface design.

Perhaps because expert systems deal with the communication of knowledge and decisions rather than simply data and information, the user interface has acquired a particular significance in the expert systems world. The CEGB is pursuing a number of expert systems projects and addressing the user interface implications of providing designers, engineers and operators with knowledge-based systems.

A major project still in its early stages is an expert system for alarm handling and fault diagnosis. The expert system is intended to be an assistant to the grid control engineers who control the transmission system at the area (i.e. regional) level. When a fault occurs on the grid, a sequence of events will take place as the grid components respond; the aim is for the system to analyse the incoming signals and determine the nature and location of the initiating fault.

In terms of the user interface for the system, the aim is to display the required information in a manner consistent with the working practices of the users. For instance, the region of the network which is the responsibility of the grid engineer is displayed on a wall diagram. Current thoughts for the user interface include displaying a similar schematic on the computer screen, allowing the engineer to select parts of the network for further study by pointing with a mouse. Also, finding the correct level of detail for information presented to the user is considered very important. One of the problems is the sheer volume of information which may arrive at the control centre; the analysis of these signals is complicated by the fact that they arrive in clusters over a period of time. At this stage, it is anticipated that the interface will provide a number of levels of information with varying degrees of detail, the first level being a simple message.

The early development is being performed using the object-oriented environment SMALLTALK-80 on a SUN workstation. The SMALLTALK-80 system makes a versatile graphics facility available to the system developer, and the combined system can also support some user interface prototyping activity.

A model for the user interface has here been immediately suggested by the working practices of the grid engineers, i.e. the prospective users. This can be contrasted with another CEGB project concerning the computerisation of a procedure for assessing structures under dynamic conditions. This procedure is contained in a document called the H0001 Report. In this project, an understanding of the prospective user activity was dependent upon the way in which the knowledge-based component developed; there were initially no precise descriptions of how the computerised version would make demands on the user.

For this reason, the early stages of the project focussed on the task of encoding the procedure in a knowledge-based form. Because the assessment procedure required access to large modelling programs, the decision was made to use the ESE/VM tool on an IBM mainframe. The anticipated requirement for diagrammatic graphics could not be met by ESE/VM itself, but such graphics were available via the use of external routines. This route, however, had limitations, and subsequently it transpired that the way the external routines were used was less than ideal for the presentation of the graphical screens required.

Part of the overall project involved the computerisation of the flow induced vibration procedure. As work on this proceeded, the limitations of the graphical presentation facilities and the response time from the mainframe (being accessed remotely) became progressively more evident. At this point, the developer of this module decided to prototype the system using a PC based expert system shell. This shell provided an improved response and, using the integrated graphics, a different appearance. This gave a different perspective on the interface requirements and provoked a more informed discussion.

At the present time, the PC version has been re-implemented using the ESE/VM tool, but the developers are now taking a wider perspective and considering target machines other than mainframes. The wider message is that only through the development of early systems (whether or not they were termed 'prototypes') could the interface requirements for this end product begin to be discussed sensibly.

This last project also illustrates how the choice of software product can place restrictions on the system developer. The following discussion concerning three welding-related systems explains why the need for good presentation capabilities resulted in a programming language being used for the interface in preference to a commercial product.

The CEGB's Marchwood Engineering Laboratory has been involved with a number of projects relating to welding technology. There are three systems aimed at providing assistance to welding engineers:

1. the selection of a welding process for stainless steel;
2. the choice of welding material when lamellar tearing is a risk;
3. the production of a welding procedure (for a welder to use directly) for CrMoV steels.

Unlike the alarm handling project where the real-time aspect must be considered in the user interface design, these welding advisors are driven by the user in a consultation-style session. Such interfaces differ from those for plant operators, for example, in that the user is an expert who needs to be given confidence in the capabilities of the system. This means that the information tends to be more detailed in nature, and also the user is given more intermediate indications as the session proceeds.

These welding advisers are PC based systems and to present the information in the desired manner it was considered necessary to create hand crafted interfaces. This was partly influenced by experiences of early PC based expert system shells which had only very limited potential for customising the appearance of the user interface. Just as there is a technological perspective on expert systems (with shells, toolkits, environments and AI languages available) so there is an MMI technological perspective, concerned with a number of different routes to the efficient and flexible production of user (and other) interfaces. This paper has already mentioned base-level languages, shells and the SMALLTALK-80 environment; another route will be discussed in the next Section.

Various points emerge from the above project discussions. Current working practices of the prospective users need to be considered in the design of the interface, as reflected in the interface work for the power system alarm handling system. It is essential that the profile (e.g. cognitive style) of the prospective user and the role of the system are properly understood so the interface can be tailored accordingly. The welding advisory systems have to provide detailed explanations to the expert user, whereas brief and clear advice is seen as necessary for a plant operator's user interface.

The nature of the information contained in the underlying system must be considered in the interface design. The construction of an early system may be necessary to bring out the interface issues. The flow induced vibration procedure interface issues were simply not accessible before the structure of the knowledge in the system had been uncovered.

One conclusion which does emerge from all the projects discussed is simply that consideration of user interface issues *is* important. Further, the important issue is to identify those features to make the interface appropriate to the users and the system.

A USER FRIENDLY INTERFACE FOR THE R6 DEFECT ASSESSMENT PROCEDURE

Background

The aim of this project - the R6 Interface Project - is to provide a user friendly interface for a program which assesses structures containing fracture mechanical defects. This assessment program is referred to as the R6 Program to distinguish it from the R6 Interface. There are some similarities between this work and the structural dynamics work described above, although the techniques ultimately employed are quite different.

The R6 Program was first made available to users several years ago. Since then it has undergone development work to enable it to be more accessible to a wider range of users. Because of the large user base, both within the CEGB and elsewhere, there are good economic reasons for making the R6 program as accessible as possible. The R6 Interface Project was instigated in order to provide an improved user interface to the R6 Program.

There are several reasons why using the R6 program directly is a non-trivial task.

1. The assessment performed by the R6 program requires significant domain knowledge to be done properly.
2. The amount of data required to do an assessment can be very considerable.
3. The type of data required by the R6 program can vary markedly between assessments.
4. The supplied data has to be correctly formatted.

A good user interface can address points (2) to (4) above, which concern knowledge about the R6 Program. The aim is not, however, to de-skill the task of performing an assessment, which will still be undertaken by a competent fracture mechanical engineer.

Two separate parts of the CEGB Research Division are involved in the R6 Interface Project. The expertise involving the underlying application program is supplied by the Fracture Section, with the design and construction of the interface being done by the Mathematics and Computing Section.

In the design and construction of the interface, techniques were taken from many areas of computing, relying quite considerably on expert systems technology. Without wishing to get caught in the trap of debating what constitutes an expert system, it is *not* claimed that the R6 Interface is an expert system. It does, however, contain sufficient aspects relevant to expert systems to merit its discussion in this paper.

The R6 Interface is, quite simply, a pre-processor. The R6 program cannot be run until all the necessary data has been supplied. Therefore, the role of the interface is to collect this data from the user.

This is not meant to imply that techniques described here are unsuitable for more tightly bound interfaces. In the case of a pre-processor, deciding which piece of data to gather next depends on the data already assembled. For an interface which is intertwined with the application program this decision may involve interaction with the application program. The difference between the two types of interfaces is only in the complexity of the decision process. Other aspects, for instance the ergonomic ones, are in principle identical.

The R6 Interface Project has been running for about a year, and still has over a year before an implemented interface goes on general release.

Project Objectives

Two key objectives affected the whole course of the Project, and both were concerned with achieving and maintaining good working relations with the client. The first was to ensure the client always felt involved in the project. This was not simply a courtesy, but a necessity since continuous client involvement was vital to the success of the project. Secondly, it was considered important to make all aspects of the work as visible as possible to the client.

A good working relation with the client was important since a learning process had to be undergone by both developers and clients alike. None of the participants had previous experience of an interface project. Because of this inexperience, the visibility objective existed in an attempt to maintain progress in the right direction.

Requirements

Some of the more general project requirements are outlined here, because they dictated the final choice of the design approach. It is the design techniques which are primarily of interest, but these requirements show what lay behind their choice.

The R6 Interface must gather a complete set of data from the user for submission to the R6 program. However, this data collection process must be made as painless as possible. This is not simply for aesthetic reasons, but because a well-designed and user-friendly interface will increase the effectiveness with which the R6 Program is used.

The visibility objective discussed above applied to all aspects of the work. This included making the interface structure comprehensible to all project participants. In other words, it was required that all aspects of the interface work should be clear, including design, documentation and code.

As fracture mechanics is an evolving subject, the R6 Program can reasonably be anticipated to undergo maintenance and enhancement during its lifetime. For this reason, the interface must be made easily extensible to allow improvements in the underlying application program to be accessed by the user.

Prototyping

This section describes the use of prototyping as a way of achieving the project objectives. Prototyping was used throughout the R6 Interface Project as an interface development approach. Its use was motivated by several factors. The objective to make progress visible could be satisfied by building and demonstrating prototypes. Similarly, client involvement could be increased through demonstrations of prototypes and discussions about their features.

At the project outset there was no clear idea of what constituted an appropriate user interface for the given application program. Demonstrating prototypes provided a method for experimentation without excessive work being necessary. Also, to make an acceptable interface, it was important to get an appropriate look and feel. This involved capturing subjective views held by the people representing the prospective users. Prototyping was seen as a way to elicit such opinions, by demonstrating a prototype and inviting comments. These opinions were incorporated in further prototypes to assess their effectiveness.

The following sections describe, in turn, a design method used to support this prototyping approach and the techniques used to implement the design.

An Object Oriented Approach to the Design

A Model of the Dialogue. This section shows how object-oriented ideas (1,2) were used to reveal the underlying structure of the R6 Interface dialogue. This is not intended as a discussion on the merits of object-oriented design in general. Rather, it is intended to show the use of object-oriented ideas use in the R6 Interface Project and to assess their impact. Briefly, object-oriented design involves studying the system by considering the objects which make up the system and the ways they interact. By grouping objects together which possess common features, computer model of the structure of the proposed system can be built up.

In the R6 Interface Project, the clients were the domain experts. The interface structure was revealed in terms of objects and their connections by a series of informal interviews. The structure found was an extremely simple one and is best summarised in the following hierarchy. Notice that the following structure makes no mention of R6: it simply describes a type of data collection system.

A session takes the form of an ***interview***.

The interview is composed of ***themes*** asked when appropriate.

Each theme consists of a collection of ***questions*** which it can put to the user.

A ***question*** takes the form of a

probe where the user submits a few answers
menu where the user makes a selection
table where the user enters data in tabular form.

This formed the basic structure. There were other objects identified, e.g. checker questions used to check the user's data. The inexperience of the Project members concerning man-machine interface issues suggested that an attempt at establishing all the system objects at the outset would have required excessive effort. Since prototyping methods were to be used to refine the system specification, this was not felt to be a serious deficiency.

<u>Effectiveness of the Object Oriented Approach.</u> Analysing the proposed interface in terms of its constituent objects together with their interactions gave rise to a very clear and simple structure, in line with the visibility requirement. Certainly the finished interface may be complex due to its size, for example, but the underlying structure is clear and concise.

There are several advantages in having such a clear structure.

1. The structure was understood by all members of the project. This improved the likelihood of detecting mistakes or irregularities in the early stages of the project.
2. An interface structure which was accessible to the R6 experts allowed them to see that the correct problem was being addressed by the interface. An obscure structure would not have inspired this confidence.
3. In terms of quality control, the more of the system the client can understand the better.

In this Project, the object-oriented design produced a highly extensible structure. For example, different question types can be added, or different types of theme. This allows new facilities to be included with only minimal disruption to the existing interface, since objects can be made to interact at a very simple level.

The object-oriented approach fitted very naturally to the task in hand, that of making a user interface. Modelling the interaction of the system with the user as an interview gave a very flexible framework. The hierarchy of objects each of which can work on the gathered data to decide whether or not they should be asked also provides a very general framework, not restricted to the specific R6 case. As mentioned above, the structure is appropriate to a more general type of data collection system.

<u>The Tool Approach</u>

<u>A Description of the Approach.</u> The name "tool approach" comes from the way the executable software is created. There are two separate components to the tool approach, the ***description***, containing all the domain knowledge, and the ***tool set*** which is the set of software tools which act on the description. (3) presents a broader discussion of software tools.

The two parts of the tool approach can be described as follows:

description - holds all the domain knowledge (cf. the knowledge base in an expert system)

- made to preserve the object-oriented structure found for the system
- can be easily extended, both in terms of having an easily extensible description language and in adding additional objects
- contains details on the appearance of the objects
- forms a readable and definitive description about the performance of the interface;

tool set - set of software tools which, in the manner of a compiler, act on the description to create an executable system (cf. the inference engine in an expert system)

- preserves the object oriented structure found for the system
- contains default settings for various appearance attributes.

The description is expressed in a purpose-built language. In the R6 Interface Project, the description language provides frame-like descriptions of the system objects. This was found to be sufficiently extensible.

It is a useful shorthand to think of the tool set as a compiler. Compilers usually work on rather general computer languages, whereas the description language in the tool approach is tuned to the task in hand.

To describe how the tool approach can be used, consider the following example which describes the creation of a particular "question" object. One of the commonest types of question required for the R6 interface is the probe, used when asking the user to supply some values.

The various parts of a probe can be summarised as follows:

requirements - to specify the conditions necessary for the probe to be asked

question - to put to the user

prompts - to specify where each required value must be entered

reply - to determine the response from the probe.

Each of these is contained in the part of the description relating to a probe, i.e. the ***probe plan***. These form the technical content of a probe, but it is necessary to get details about the appearance of a probe as well. This can be done by prototyping a probe and inviting comments. It is necessary to have some tools which convert this probe plan into an executable probe object. Such tools include, for example, screen handling tools for putting text on the screen with specified colours, font, and size. The tools are then applied to the plan to make the executable probe object. This executable probe can then be demonstrated to the people who represent the prospective users of the system. Changing the appearance can be done by altering the probe plan and re-applying the tools. This can be repeated until the appearance is deemed acceptable.

Such prototyping can be used for all the objects which appear to the user in order to elicit the required appearance details. Similarly, the prototyped objects can be linked up to form a more extensive prototype. This can then be demonstrated to assess the feel of the system, and again can be altered considerably by simply changing the description.

The description part of the tool approach forms a very useful part of the system documentation. This is not a claim that the tool approach is self-documenting since, for instance, the description contains no information about the solution strategy. However, the description does provide a precise and readable record of the domain information contained in the system.

This is principally throw-away prototyping of the description and incremental prototyping of the tools. Once extensions have been made to the description language and the tool set to admit a new object type, the creation of instances of an object is trivial. Objects may be added to a prototype by adding plans for those objects to the description. This does not involve any programming language code and can be done by someone not versed in the language used for the software tools. The description language is designed to be concise, so only the absolutely essential information is needed.

Effectiveness of the Tool Approach. The benefits brought to the Project by the tool approach are concerned largely with human issues. In terms of interacting with the clients, the use of rapid prototyping and frequent demonstrations was extremely successful. The demonstrations were largely responsible for the good relations with the clients during the Project. They felt involved throughout and could see good progress being made. Also, the prototypes proved an excellent way to elicit the subjective details about the look and feel of the interface.

In terms of the R6 Interface Project, the themes and their constituent questions were constructed from a specification supplied by the R6 experts. Once this specification has been available, the average time to construct an R6 Interface theme has been one week. This includes creating the theme description, applying the tool set and testing the resultant executable theme. Given that all the R6 detail in the finished interface will be contained in about eight themes, it is clear that the tool approach offers some real benefits. Of course, it takes time for the domain experts to create the initial specification which gets turned into a theme description, but this is time spent considering how to build the interface rather than how to beat the computer system.

A frame-like representation for the basic plans of each object makes the description language easily extensible. This was particularly important in the R6 Interface Project because the specification for the system was incrementally refined rather than defined at the outset.

The tool set was also made extensible so that new additions to the description language could be compiled. This is described in the next section on the use of Functional Oriented Design.

To summarise, the tool approach was found to be very effective in the R6 Interface Project for the following reasons.

1. It allowed the implementation of the object-oriented structure of the interface.
2. It enabled rapid prototyping to be performed which was both popular with the R6 experts and which allowed the appearance of the system to be customised.
3. It enabled fast development, with important contributions by people who had no knowledge of the tools' programming language.
4. The description part of the tool approach serves as a readable and precise guide to the behaviour of the interface.

Functional Oriented Design

Description. The term 'functional oriented design' is meant to parallel that of object-oriented design. Functional oriented design is simply a way of viewing everything as a function. Functional programming (4) emerges from functional oriented design in the same way that object-oriented programming stems from object-oriented design.

In a functional oriented design, the overall problem is addressed using a functional decomposition approach. One difference between functional oriented design and more traditional software design is that the idea of the system state is not present in the functional design. The important constraint imposed by being strictly functional is that functions return values *without causing any side effects*.

Functional oriented design was used in building the R6 software tools. Since the action of the tool set is to convert the description into executable code, the tool set can therefore be considered as a function which performs this mapping.

Effectiveness of Functional Oriented Design. One of the features of using functional oriented design is that the resulting software is highly structured. Considering R6 again, the tool set showed a very clear breakdown of the compilation task it had to perform. This helped conceptually as well as in the implementation, because none of the functions written had to solve difficult tasks. The no-side-effects constraint imposed by the functional approach made it extremely difficult to create large, unwieldy functions. The functional ideas therefore forced the software tools to be small and manageable.

The functional tools which formed the R6 tool set were much easier to test and debug than if, in some fashion, the tools had operated on a system state.

In the R6 Interface Project, the network of functions comprising the tool set was printed out automatically providing a very useful part of the documentation. This was especially useful in the testing phases.

One significant drawback with functional oriented design did emerge during the Project. Although the functions themselves were simple, the sheer number of them became rather intimidating. This conceptual overload was addressed in various ways.

1. The network of functions was generated automatically by a function to analyse the tool set.
2. The facility for arbitrarily long function names meant the names could be chosen to reflect the purpose of the tool. The network was therefore useful in summarising the relationships between the tools.
3. The problem to build the compiler, i.e. the tool set, was decomposed so that this network of functions did not have a uniform connectivity. The network consisted of regions of high connectivity with relatively few links between the regions. This meant the individual clusters could be treated in relative isolation thus reducing the scale of the conceptual problem.
4. Every function was documented, including details on where it fitted into the overall tool set as well as how it operated.

Current status of the interface product and the toolkit

The R6 Interface Project still has over a year to run before an implemented interface goes on general release. However, the prototype interfaces built so far have been demonstrated to a number of interested parties and have been well received. It is not expected that any of the subsequent refinements will render any of the above conclusions invalid.

The tool set potentially has much wider application than to the R6 Interface and over the next year we will be looking for opportunities to use both the tools and the ideas embodied in their construction on further interface projects.

CONCLUSIONS

The main conclusion to come from work done within the CEGB on user interface issues to identify the ***appropriate*** interface facilities for the finished system.

The discussion of a selection of CEGB projects also indicates some of the factors to be considered when determining the appropriate interface facilities. These factors are itemised below.

1. The prospective users must be considered, both in terms of their working practices as well as their skills.
2. The role of the interface and the environment in which the system is to be used are both important to the interface design.
3. The structure of the knowledge in the underlying system must be taken into account in the interface design.

Conclusions arising from the R6 Interface Project discussion can be drawn on two different levels.

1. In project management terms, an active policy to keep all aspects of the work visible to all the project members can help achieve a good relationship with the client.
2. Concerning the interface design, the combination of techniques described can enable an appropriate interface to be produced using prototyping to refine the interface specification.

ACKNOWLEDGEMENT

The work was carried out at the Central Electricity Research Laboratories and is published by permission of the Central Electricity Generating Board.

REFERENCES

1. R.S. Pressman. Software Engineering: A Practitioner's Approach, 2nd ed., McGraw-Hill International Editions, 1987.
2. B.J. Cox. Object Oriented Programming: An Evolutionary Approach, Addison-Wesley, 1987.
3. D. Ince, M. Woodman, S. Hekmatpour. "The application of some artificial intelligence tools and techniques in software engineering." Software Engineering: the decade of change, IEE Computing Series 8, Peter Peregrinus Ltd., London, 1986.
4. H. Glaser, C. Hankin, D. Till Principles of Functional Programming, Prentice-Hall, 1984.

Intelligent Interfaces to Expert Systems Illustrated by a Programmable Signal Validation System

BJORN FROGNER
Expert-EASE Systems, Inc.
1301 Shoreway Road
Belmont, California 94002, USA

ABSTRACT

This paper discusses a software tool for the development of effective interfaces to an expert system. These are interfaces to end-users, application developers, as well as interfaces to other software modules. The application of this tool is illustrated by discussing a "programmable" signal validation capability. The objective of this discussion is to demonstrate how easily an expert system application can be configured through the use of graphics to reflect changes in instrumentation, plant configuration or signal validation logic.

PROBLEM DESCRIPTION

In broad terms, the current methods for signal validation can be divided into the following categories [1,2,3,4,5]:

- *Reasonableness checks.* Complete failures typically result in high or low readings; i.e., at the extreme ends of the scale. Such failures can easily be recognized by checking if the measured values are within the expected bounds.
- *Majority vote.* In those areas where there are three or more redundant readings, a relatively straightforward majority (e.g., 2-out-of-3) vote can be used.
- *Consistency checks.* There are several areas where there are different but dependent variables (e.g., the pressure at different points in a steam line) that are known to have very close relationships. Such measurements can easily be checked for consistency.
- *Rate-of-change.* By knowing the physical processes, one can determine how fast a detector reading can be expected to change and then classify changes that are significantly faster as being unreasonable; i.e., due to malfunctions in the instrumentation or the electronics. A wide range of sophistication exists in this area; from fixed thresholds on rate-of-change of individual measurements to multivariate statistical models [6].

- *Analytical models.* The use of models for analytically derived "measurements" or in conjunction with state estimators can result in high diagnostic sensitivity across a wide range of operating conditions [5,7,8].
- *Parity space.* This approach [1] presents a common metric for handling analytical redundancies that involve variables of different kinds; e.g., pressure and temperature.
- *Expert systems.* This technology has only recently been investigated [4,6,8,9] in the context of signal validation and only limited experience exists yet as to its exact contribution in this area. The expectations are that it can integrate all the methods presented above and additional features (e.g., complex heuristic experience) can be incorporated. This is the major focus of this paper.

The software tools presented in this paper can be used to implement all the methods described above in an integrated manner.

SOFTWARE LAYERS

To design effective interfaces to expert systems, it is helpful to review the relationship between expert system shells and other programming environments. Figure 1 illustrates the various levels of software tools from the operating system (OS) as the innermost layer to the application code as the outer layer.

- The *operating system* consists of very low level languages that almost never is dealt with by the application developer nor the end-user.
- The *programming level* consists of standard programming languages (e.g., C, Fortran, Lisp), communications software, window screen managers (e.g., X-Windows, Presentation Manager), etc. Development at this level results in software that is fairly easy to port to other computers. Furthermore, there is a substantial flexibility in the functionality. However, development at this level typically involves large cost.

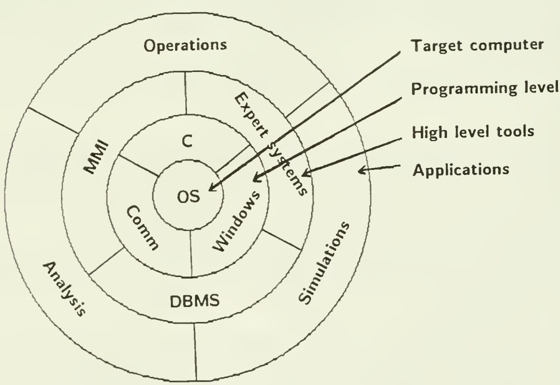


Figure 1. Overview of Software Layers Involved in Development of End-User Applications

- The *tool level* consists of generic high level tools such as Data Base Management Systems (DBMS), Expert Systems (ES), Man-Machine Interface (MMI) packages, etc. The objective of the tools at this layer is to elevate the application developer to a higher level to improve the productivity of development. Furthermore, if the right tools are used, a high degree of portability between computers can be achieved.
- The *application level* consists of the applications code which computes, analyzes, or otherwise performs the job that is of interest to the end-user. If the application code has utilized effective tools, its portability, maintainability and flexibility will be substantially enhanced.

One highly effective way of improving the productivity of application development is to increase the functionality, standardization and integration of the software at the "tool level." This is the underlying motivation for the work described in this paper.

REQUIREMENTS FOR INTEGRATION OF EXPERT SYSTEMS

To effectively imbed an expert system in an integrated environment it is necessary to consider the following capabilities:

- *Easy to Use.* The interface to the expert system must be easy to learn and productive to use both for the developer and the end-user. It must be intuitive, self-guiding (internal help messages), robust to errors, rich in graphics and menu driven. It is important to realize that the end-user wants productive solutions (not technology) while the developer wants productive tools (which may include technology if it simplifies the implementation).
- *Easy to Modify.* It must be easy for the end-user to update the knowledge base (KB) as a result of changes in plant configuration, status or condition. Modification of plant configuration should be done graphically and the KB should automatically reflect these changes. One way to achieve this is to code the rules at the class level and make a strong correspondance between the objects in the expert system and the objects (icons) in the graphical environment.
- *Object-Oriented.* Both the expert system and the surrounding environment (e.g. the graphics) should preferably be object-oriented to facilitate representation of physical systems.
- *Interface to Data Base.* The expert system needs an effective interface to a data base to find the values that are needed in the reasoning. Extensive interactions with the user to determine plant conditions and other values is not acceptable.
- *Use of Models.* Causal models as opposed to "compiled" knowledge, as represented by production rules, is very desirable as an augmentation to an expert system shell. The reason for this is that in a causal model, there will be no fixed set of rules and, thereby, fixed dependencies within the system.
- *Complex Reasoning.* In a typical application, a large fraction of the rules involved are quite simple and not worthy of the complication of being processed by a sophisticated expert system shell. Thus, the expert system should be used to perform the higher level reasoning while the low level reasoning should be taken care of by simpler means; e.g., decision tables.

SOFTWARE MODULES

Block Diagram Overview of Major Modules

The signal validation system presented in this paper was developed by the integration of three existing and widely used software tools as shown in Figure 2.

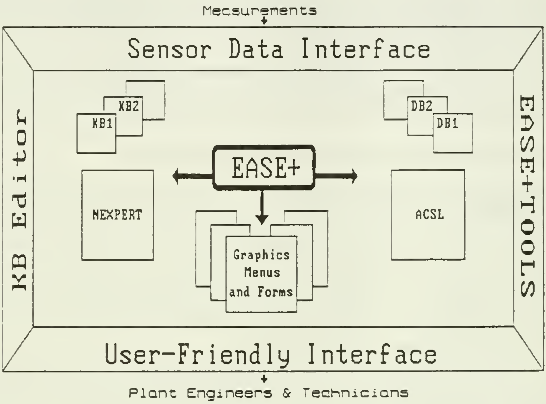


Figure 2. Overview of Major Software Modules

- *EASE+* [10] is the overall environment for integration of all the modules and it performs the interface to the end-user. It has extensive capabilities in the areas of graphics, data base and user-friendly features.
- The *NEXPERT* [11] expert system "shell" is the means of processing the domain-specific knowledge bases. NEXPERT draws on the current real-time values present in the data structures when it needs specific values from the measurements.
- The *ACSL* [12] module is used to integrate the simulation models forward in time. It can also draw from the knowledge base to determine its response to the reasoning processes.
- *KBs.* There may be any number of modular knowledge bases (KBs) supporting the expert system reasoning. These KBs contain the plant specific signal validation logic.
- *DBs.* A modular approach is also used for the data bases. These DBs contain the real-time data coming in from the sensors as well as intermediate calculational results.
- *Sensor Data Interface.* This module takes care of bringing the necessary plant information into the internal data bases.

- *User Interface.* The end-user deals with a highly effective interface that uses plant schematics (for display of instrumentation), menus (for choosing options) and forms (for data entry).
- *KB Editor.* Powerful KB editors available in NEXPERT can be used to modify the knowledge bases.
- *EASE+TOOLS.* Engineers who are qualified to modify the applications aspects of the software can use the variety of high level tools available in EASE+. These tools can be used to modify the graphics, add to the data base, integrate new analysis capabilities, etc.

EASE+ Capabilities

EASE+ [10] consists of two parts: a) a high level software tool-kit for development of specific applications and b) a runtime software module that functions as a delivery environment. Using this tool-kit in an interactive manner, a developer can create full-color dynamically updated schematic diagrams, generate the necessary data base structures, interface with external programs, implement the logic flow associated with a specific application, etc. With the EASE+ run-time module, an end-user can interface with an application through graphics, menus, and data entry forms.

In the context of the expert system, EASE+ serves as the overall operating and control environment performing the following functions:

- *Instantiation.* By the user interactively connecting predefined graphical icons (objects) on a CRT screen to reflect the configuration of the instrumentation and associated validation logic, EASE+ informs the expert system that it must instantiate the relevant objects at run-time.
- *Initiation of analysis.* Triggers execution of the expert system through user selection of an appropriate option from a menu or activated automatically upon recognition of a problem.
- *Focusing of the reasoning.* Provides an interface between the knowledge base and color schematics of the plant subsystems. These graphic representations consist of a series of interconnected icons representing individual components in the plant. The users will be able to focus the analysis on a particular subsystem or component by placing the cursor on the appropriate icon.
- *Presentation of results.* Informs the users of the results of the analysis by highlighting the affected components on a color schematic and providing a text description of the likely problems.

NEXPERT Capabilities

NEXPERT [11] is an advanced and widely used expert system shell developed by Neuron Data, Inc. The following features are important for the signal validation problem:

- *Object-oriented structure* - this feature allows structuring of the knowledge base according to the hierarchical structure common to most engineering systems.
- *Forward and backward chaining rules* - IF...THEN...ACTION type of rules to contain the signal validation logic.
- *Methods* - this feature facilitates the integration of arbitrary processing, procedures or code at almost any point in the reasoning.

- Ability to specify a *context structure* for rules - this feature allows effective control of the reasoning process.
- Ability to access *external routines* or perform other user-specified functions such as external calculations or solicit the users' responses to assist in the analysis.
- Ability to *volunteer data* to NEXPERT prior to the start of the session - this feature allows the expert system to be tied to a real-time data base that automatically supplies it with the latest information needed for the reasoning.
- Ability to *focus the reasoning* (concentration on a particular line of thought) externally by suggesting likely conclusions prior to the start of the session - this feature enhances system efficiency by allowing the user to rule out unlikely conclusions before they are considered.

ACSL Capabilities

ACSL [12] is a widely used software tool for modeling and analysis of continuous-time systems described by time-dependent, non-linear differential equations or transfer functions. Integrated underneath the EASE+ environment, ACSL enables the user to perform the following functions:

- *Model building*: Graphically construct predictive simulation models of the plant.
- *Parameterization*: Specify various parameters and options through data forms.
- *Execution*: Initiate and control the execution of the simulation models.
- *Results*: Display the results through x-versus-time plots, as numbers on graphics displays or as reports.

FUNCTIONAL DESCRIPTION

Implementation of Signal Validation Logic

Assuming that the necessary instrumentation, associated electronics, and computer processing hardware needed for driving the signal validation software are available, implementation of the signal validation software for a specific application then requires the plant personnel to go through the following steps:

- *Graphics.* Using the EASE+ tools, the user can generate graphical representations of the plant instrumentation diagrams and schematic "mimic" diagrams of the associated plant subsystems. These diagrams are used to identify graphically how the sensors are related to the plant and they are available for real-time data display as well.
- *Models.* The simulation models that are needed can be developed by using ACSL as the basic simulation language. Block diagram graphical representation of the models is available as well as direct access to the underlying programming languages (FORTRAN and C). Assigning values to the many parameters can easily be achieved by "pointing" to the appropriate iconic representations of the associated components.
- *Knowledge base.* The third step involves developing the application specific knowledge base; i.e., the logic needed to validate the sensor readings. This information is prepared by filling out "forms" using the knowledge base editors available in NEXPERT.

Coupling Between Graphics, Models and Rules

There are two types of graphical models that can to be built. The first is the graphical "mimic" representation of those parts of the plant that the user wants to monitor. The second is the ACSL simulation models for these same systems. The user can build these models by using the preestablished library of icons that are available as the basic building blocks. Beyond the pictorial appearance on a screen, the graphics has the following objectives: establishing connectivity between the physical components, instantiation of the objects in the knowledge base, representation of the hierarchical relationships and easy access to the data base.

These two graphical representations will in general have many commonalities since they relate to a different "view" of the same system. Thus, they are linked tightly underneath the user level. Since the system may consist of a hierarchical assembly of objects, it shares the "knowledge" about the individual objects regardless of whether the graphics representation is for the benefit of EASE+, ACSL or NEXPERT. Furthermore, the user can build up his graphical representation of the model by using basic ACSL type of icons (i.e., adders, multipliers, etc.) at the lower levels and then put them together as "mimic" diagram representations of the plant at the higher levels. In this manner, the graphics, modeling and knowledge base capabilities have been very tightly integrated.

Diagnostic Process

The major steps that the signal validation software performs during real-time processing are:

1. Obtain the measured data from the appropriate data acquisition system.

2. Run the simulation model one sampling interval forward in time to obtain a corresponding predicted value for each "modeled" parameter.

3a. If predicted value is available compare the measured and predicted values.

3b. If redundant measurements are available compare redundant values.

4. Use the rules in the knowledge base to determine if the differences identified in step 3 are significant and what action to take with respect to these differences.

5. Individual sensor quality tags are determined by incorporating uncertainty calculations.

6. The results of the signal validation are stored in the data base. Update displays and communicate with the user if so desired.

7. After having obtained the best composite reading, the predicted values are updated according to whatever state estimator algorithms the user has specified.

The software displays the plant system and subsystem model, presents bar-charts of measured values and the time-evolution of chosen signals. When a significant discrepancy occurs, the loop is interrupted and a menu pops up automatically for the user to review the explanations.

KNOWLEDGE REPRESENTATION

An important part of any expert system implementation is the development of a good framework for representing the knowledge that should be captured. The concerns guiding the knowledge representation are: constraints of the selected knowledge engineering software, effectiveness of implementation, ease of maintenance and usefulness of final system. The major representational schemes that are needed are:

- The object hierarchy.
- Object-oriented inheritance to effectively divide plant components into a hierarchical class-structure which simplifies assignment of component attributes.
- Production rules to express heuristic knowledge.
- Uncertainties due to errors in detector readings and incompleteness of the heuristic rules.
- Access to mathematical model calculations.
- Control structures to make "shortcuts" in lengthy reasoning sequences.

In an object-oriented expert system shell like NEXPERT, one ordinarily starts building the knowledge base by first mapping out the object structure. The object structure should follow the hierarchical structure of the particular system. One can then prepare the rules that specify the behavior and reasoning associated with these objects.

After having developed the objects and the rules, one has to control the reasoning process. This is particularly important for signal validation since processing speed is of the essence. NEXPERT is controlled by an "agenda" that determines what to check next, what information shall be passed along, etc. This agenda is controlled automatically in three ways through EASE+:

- Selected values are "volunteered" to NEXPERT and the effects are then propagated throughout the knowledge base by forward chaining.
- One or more hypotheses are "suggested" to the agenda and all the conditions attached to the associated rules are investigated to determine if the hypothesis is true. This is a backward chaining functionality. Restrictions (or focusing) of the suggested hypothesis can be set to:
  - Quit the reasoning when the hypothesis has been proven true;
  - Continue the reasoning without checking the suggested hypothesis again when it is proven true; and
  - Exhaustive firing of all the rules in the knowledge base.
- "Data propagation." Data that were generated in the action part of a rule will be propagated to other rules. Controls are available to turn such propagation on and off anytime or to restrict the effect to be either local or global.

The effective utilization of these capabilities is important for real-time applications where speed of response is of the essence.

SIMULATION MODELS

Physical Models

To provide an example that demonstrates most of the available features, a simple model of the reactor water level in a Boiling Water Reactor (BWR) was used. The essence of this model is as follows. If the input flows from the sources exceed the output flows, then the reactor water level (RWL) will go up, if the input flows are less than the output flows, then the reactor water level will go down. Furthermore, as the pressure, p, in the vessel increases it will collapse the steam bubbles, while if the pressure decreases it will cause flashing. This effect can have a significant influence on the water level during fast transients. Thus, the model was as follows:

$$d(RWL)/dt = (\text{flow in} - \text{flow out})/\text{area} + \text{constant} * dp/dt$$

The flow rates and the pressure are dependent upon other state variables. Models of this type have been demonstrated to be implemented easily by using the graphics user interface available in EASE + ACSL.

Use of Observers and Kalman Filter

In deterministic processes or processes where the noise intensities and uncertainties are small enough to be ignored, the appropriate method for filtering measurements against a dynamic plant model is the Luenberger observer [5]. In processes containing strong stochastic components, our experience indicates that the Kalman filter [7] is usually an appropriate tool. Fault detection can then be done by investigating the statistics associated with the residuals (differences between predicted and measured values). The essence of the residual-based technique is the correlation of filter optimality with failure detection. If abnormalities appear, changes in the statistical properties of the residuals are expected to occur. Therefore, by performing statistical tests on the filter residuals it is possible to determine whether or not a failure in the system has occurred.

TEST PROBLEMS

BWR Water Level Test Case

There are typically four different kinds of water level instrumentation in a BWR: narrow range, wide range, yarway and refuel-mode sensors. There are usually three narrow range sensors, two wide range sensors and two yarway sensors. This redundancy gives rise to a wide range of possible cross-comparisons as well as weighted averaging. The logic needed to evaluate such redundancy was effectively implemented by the available expert system capabilities.

Transient data from a simulator of a BWR were obtained for various significant plant transients. Each transient was a second-by-second record of the simulator's entire analog and digital data base. There were a few hundred analog parameters and several hundred digital parameters recorded each second.

When the transient began, the model of the reactor water level used a mass balance equation on water inflows and the steam outflows to compute the dynamically changing water level. A Kalman Filter was used to adapt the model to the aggregate water level reading after each sampling interval. If significant differences were detected, the data were analyzed and warnings of inconsistencies made available.

Figure 3 shows a typical CRT display. In the upper left quadrant the plant schematics appear in colors to highlight problem areas when necessary; the recent trend for a chosen sensor reading versus corresponding prediction appears in the upper right quadrant; a comparison bar chart for some selected sensors are shown in the lower left quadrant; and finally in the lower right quadrant there is a list of options available for investigating this problem.

Turbine-Generator Test Case

To exercise the signal validation concepts with respect to real-time monitoring of actual sensor readings, a demonstration system capable of monitoring and evaluating a limited portion of the Balance-of-Plant system for a nuclear power plant was developed. The software, sensors and electronics that were put together were used to evaluate real-time changes of operating parameters (e.g., thrust-bearing wear rates) with normal wear rates experienced by equipment with similar characteristics. Bearing temperature, generator hydrogen makeup flow, thrust bearing wear, shaft vibration, and lubricating oil quality were the operating parameters and conditions chosen for evaluation. Figure 4 shows a schematic representation for this system.

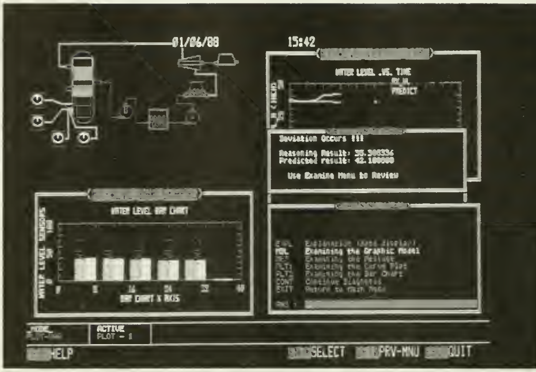


Figure 3. Typical Display Produced During the Water Level Test

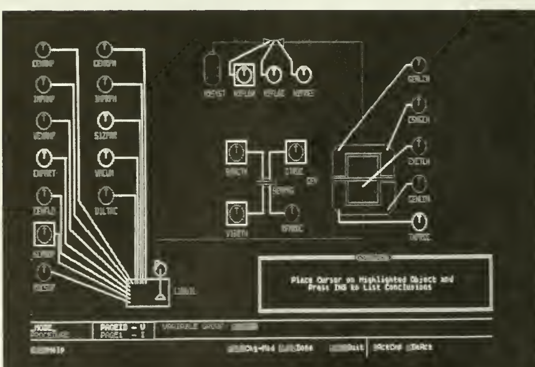


Figure 4. Instrumentation Schematics for the Turbine-Generator

At the end of the diagnosis, the expert system reported to the user its results in the form of , conclusions and recommendations. If sufficient data did not exist in the knowledge base to form definite conclusions or make definite diagnoses, requests for additional input from the user were made. In addition, if an "alarm" flag had been set for a parameter, the user could be notified along with a recommended action. This recommended action was dependent upon the state of other operating parameters and information possessed by the expert system.

Data Acquisition

The signal generation and data acquisition system used in this test were developed by Volumetrics, Inc. The hardware needed to build the system was relatively simple and it used readily available instrumentation and electronics. To actually implement a similar system in a power plant would require minimal modifications to existing plant equipment. In many cases, existing plant instrumentation and computers can be utilized. The signal generator box for this demonstration consisted of a micro-processor controlled "black box" which had a readout panel for reading the current value of each of the programmed parameters.

The output from the signal generation box consisted of an RS-232-C channel which periodically sent out an ASCII coded message. The signal values were repeated every two seconds. The values were controlled by control knobs. By choosing the various combinations of outputs, the software could be made to exercise most of its logic reasoning processes.

Validation of Key Sensor Inputs

In this test, the processes and signals being monitored could not be simulated conveniently using physical models. Thus, signal validation was accomplished by checking the sensor readings for reasonableness and consistency with other physically related signals. This reasonableness/consistency checking approach to signal validation was implemented easily by using the expert system. The only real complication in the process was in the determination of which signals needed to be validated and which other signals should be used to support this validation process. Unless proper care was taken in selecting these signals, consistency checking could become a circular process in which multiple signals were being validated simultaneously by comparing them to each other.

To illustrate these issues, consider the hydrogen cooling subsystem of the turbine-generator system. For the hydrogen subsystem, the most important indicator of a potential subsystem malfunction is the hydrogen flow rate. When the hydrogen cooling subsystem is functioning normally, hydrogen is supplied to the generator at a steady rate of 45 SCFD (standard cubic feet per day). Any variation in this makeup flow rate is indicative of a potential problem. The diagnostic knowledge base for the hydrogen cooling subsystem therefore treats hydrogen makeup flow not equal to 45 SCFD as a necessary condition for all subsystem problems. When this condition was met, the knowledge base evaluated a variety of other signals (e.g., hydrogen flow rate-of-change, hydrogen line pressure, hydrogen concentration at various locations in and around the generator) to identify the most likely source of the problem and recommended appropriate corrective action.

This hierarchical approach to the diagnostic process indicates that the hydrogen flow measurement is the key to proper functioning of the monitoring system and should therefore be subjected to routine signal validation. The remaining signals were then used as consistency checks to perform this validation in the following manner:

1. If the measured hydrogen makeup flow is less than 45 SCFD: Hydrogen line pressure and the rate-of-change of hydrogen flow are checked for indication of depletion of the hydrogen supply bottles. If both of these indications are normal, then the hydrogen flow measurement is assumed to be invalid.

2. If the measured hydrogen makeup flow is greater than 45 SCFD: Hydrogen concentration around the generator and the rate-of-change of hydrogen flow are checked for indication of a hydrogen leak. If both of these indications are normal, then the hydrogen flow measurement is assumed to be invalid.

3. If the measured hydrogen makeup flow is equal to 45 SCFD: Hydrogen line pressure, rate-of-change of hydrogen flow and hydrogen concentration are checked. If two of these indications are abnormal and consistent with each other, then the hydrogen flow measurement is assumed to be invalid.

Expert System Actuation and Results Display

For the current prototype, the expert system diagnosis was actuated manually via a menu selection or automatically as the real-time data were received based upon the current value of three key indicators of generator system trouble. These key indicators (hydrogen flow rate, bearing temperature rate-of-change and lube oil screen differential pressure rate-of-change) were checked for any indication of potential problems and, if any of the three were outside of their normal range, the signal validation analysis was actuated. Once actuated, it first performed a validity check on the three key indicators as described above. If the abnormal indication was invalid, the session was terminated and the invalid input was flagged to the user. If the abnormal signal was valid, or if a normal indication was found to be invalid, the expert system checked the remaining analog and digital signals to determine the most likely problem. When the diagnostic session was completed, the results of the diagnosis were displayed graphically in the following manner:

- If a problem was detected, the icon associated with the problem was highlighted in red. Icons representing support components that were functioning normally were displayed in green.
- For each identified problem, the "dials" representing the analog signals whose values were indicative of that problem were highlighted in yellow. "Dials" representing analog signals whose values were normal or otherwise unrelated to any identified problem were displayed in green.
- If any key signals were found to be invalid, the "dials" representing these signals were highlighted in red.

To obtain a text description of the identified problems, the user could position the cursor on the appropriate icon or "dial." As shown in Figure 5, this text description identified the bad signals and the reasoning behind these results.

SUMMARY & CONCLUSIONS

EASE+ has been used as an integrating environment in many applications and with many codes. The EASE+NEXPERT combination has been demonstrated particularly viable and the integration with ACSL has proven potentially very powerful. The integration of EASE+, NEXPERT and ACSL has been evaluated for signal validation in two tests:

- Validation of the signals for the reactor water level in a Boiling Water Reactor (BWR) using high quality data from a training simulator. A representative knowledge base, a simple mass-balance model, approximate sensor noise and a reasonably realistic simulation scenario have been implemented and successfully demonstrated.

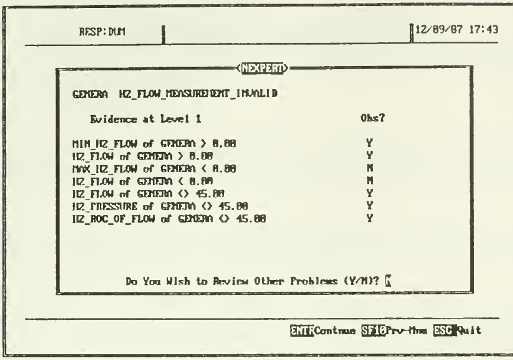


Figure 5. Identification of Suspect Signals and Associated Explanation

- Validation of the signals for a simulated turbine-generator diagnostic system in a nuclear power plant using a signal generator data acquisition system developed specifically for this project. This demonstration successfully tested the use of actual real-time signals.

The major benefits from using an expert system approach compared to conventional programming languages for signal validation are:

- *Representation.* NEXPERT is rich in its ability to represent complex problems. For example, the object-oriented capabilities represent a natural and powerful means of representing hierarchical systems, subsystems and components. Most of the signal validation logic that is needed can readily fit into the NEXPERT knowledge base.

- *Modifications.* Since the knowledge base is separate from the general code, it is easy to modify. This is very attractive since much of the logic associated with signal validation is application/plant specific and it needs occassional update.

- *Explanations.* The expert system is able to explain its line of reasoning; i.e., supply the pieces of information behind a conclusion. This is important for building confidence in the results. Explanations can also be programmed into systems implemented in conventional languages; however, an extra effort has to be put in to get that benefit.

Through the integration of EASE+, NEXPERT and ACSL, these capabilities are now available to a wide class of users through high level interactive tools instead of requiring extensive programming and knowledge engineering training.

ACKNOWLEDGMENTS

The signal validation aspects of the work reported in this paper was funded by the U.S. Department of Energy under contract DE-AC03-87ER80539.

REFERENCES

1. C.H. Meijer and J.P Pasquenza, *On-Line Power Plant Signal Validation Technique Utilizing Parity-Space Representation and Analytic Redundancy*, EPRI Final Report NP-2110, Nov. 1981.

2. B.J. Benedict, S.S. Godbole and R.I. Lutz, *Validation and Integration of Critical PWR Signals for Safety Parameter Display Systems*, EPRI Interim Report NP-4566, May 1986.

3. E. Kujawski, et.al., "Signal Validation Software Utilizing a Generalized Decision Estimator for Common Cause Failures," *ANS 1985 Topical Mtg on Computer Applications for Nuclear Power Plant Operation and Control*, Pasco Washington, Sept. 1985.

4. B. Chandrasekaran, "An Artificial Intelligence Approach to Sensor Conflict Detection," ibid.

5. L.G. Niccoli, N.P. Wilburn and R.W. Colley, "Detection of Instrument or Component Failures in a Nuclear Plant by Luenberger Observers," ibid.

6. B.R. Upadhyaya, T.W. Kerlin, O. Glockler, Z. Frei, L. Qualls and V. Morgenstern, "An Integrated Approach for Signal Validation in Nuclear Power Plants," *ANS Topical Meeting on AI and Other Innovative Computer Applications in the Nuclear Industry*, Snowbird, Utah, Aug 31 - Sept 2, 1987.

7. B. Frogner, B. Friedlander and H.S. Rao, "Methods for Identification of Dynamic Systems," *Nuclear Science and Engineering*, Vol. 64, pp. 644-656, Oct. 1977.

8. J.A. Hassberger, *Simulation-Based Expert System for Nuclear Power Plant Diagnostics*, PhD Dissertation, Department of Nucelar Engineering, University of Michigan, Oct. 1986.

9. S. Hashemi et. al., "An Expert System for sensor Data Validation and Malfunction Detection," *ANS Topical Meeting on AI and Other Innovative Computer Applications in the Nuclear Industry*, Snowbird, Utah, Aug 31 - Sept 2, 1987.

10. EASE + Technical Description, Version 3.0, Expert-EASE Systems, Inc., June 1989.

11. NEXPERT Technical Description, Produced by Neuron Data, Inc., 1989.

12. "Advanced Continuous Simulation Language (ACSL)," *User Guide/Reference Manual*, Mitchell and Gauthier, Associates, 1981.

NUCLEAR POWER PLANT APPLICATIONS

The Utilization of Expert Systems within the Nuclear Industry

J. A. BERNARD and T. WASHIO
Nuclear Reactor Laboratory
Massachusetts Institute of Technology
138 Albany Street
Cambridge, Massachusetts 02139, USA

ABSTRACT

The utilization of expert systems within the nuclear industry is examined. Topics reviewed include factors motivating the industry to develop expert systems, areas of application, and issues related to acceptance. It was found that expert systems, as currently conceived, can be used for managerial tasks such as ensuring regulatory compliance and for interactive diagnostics. However, it is unclear that the technology can be utilized for real-time diagnostics and guidance. For this to happen there must be substantial improvements in the man-machine interface and extensive experimental assessments of the technology.

INTRODUCTION

This paper examines the utilization of expert systems within the nuclear industry. It is a state-of-the-art review that draws heavily, but not exclusively, on a book that the authors recently completed on this topic [1]. Some 287 expert systems are identified in that book as either under development or in use within the nuclear and commercial electric power industries. One of the book's more important contributions is that it places this activity in perspective. Major areas of application are identified. These include systems for use as engineering tools, the capturing of human expertise, plant design, facility management, maintenance planning, interactive diagnostics, real-time diagnostics, decision support, emergency response, cognitive models, and control. Each application is assessed in general terms relative to the capabilities of the technology. Specific systems are then described. The result is that the strengths and weaknesses of the expert systems approach become apparent. In addition to delineating areas of application, the book also discusses the motivation of the nuclear industry for developing expert systems and factors relevant to the successful implementation of those systems. Included as part of the latter topic are criteria for problem selection, observations on the characteristics of successful nuclear expert systems, a discussion of operator needs and the man-machine interface, and an overview of regulatory perspectives. The book concludes with a section on 'lessons learned' and suggestions for enhancing the prospects for the successful implementation of nuclear expert systems.

The specific objective of this paper is to provide a concise summary of certain portions of the aforementioned book. The areas selected for presentation are (1)

This is reprint of a paper presented at the 1989 American Control Conference and published through the American Automatic Control Council or AACC.

clauses. Another benefit that accrues to the nuclear industry from this explanatory aspect of expert systems is that it facilitates the preparation of the written justifications that must be maintained as documentation for most decisions, even routine ones.

A third major advantage to the usage of expert systems within the nuclear industry is that much tedious work can be eliminated. For example, checking planned maintenance and scheduling activities against the applicable quality assurance standards and surveillance requirements is a process that is normally performed by skilled, experienced personnel. Individuals with less training might not be capable of differentiating rules that are appropriate from those that are not. Hence, such tasks are often a heavy burden on the most talented individuals. An expert system can do much of the drudgery and leave skilled personnel free to address those few questions that really merit their attention.

Areas of Application

Some 287 expert systems are identified and discussed in the actual book. These are summarized by topic and national origin in Table One. The categories to which the individual systems have been assigned were chosen so that there would be a logical progression from the more traditional applications of expert systems to some of the more esoteric uses to which the technology is being applied within the nuclear industry. Such a classification scheme is, of course, superficial because disparate applications are being attempted in parallel rather than in a serial fashion. Also, a given system may combine both basic and advanced concepts. Nevertheless, such an ordering is useful because it focuses on trends and reveals unresolved issues. Among the findings of the study are that:

- Expert systems are most readily developed and implemented if those responsible are cognizant of both the technology in question and A/I techniques. Given that it takes years of study and experience to master any field of engineering, it is far more practical for an industry specialist to learn and apply the methodology for constructing an expert system than for an A/I practitioner to acquire a thorough knowledge of the industry. Accordingly, the electric utilities should continue to provide opportunities for their engineering staffs to learn about expert systems technology. Also, they should be pressing for the inclusion of courses on expert systems in university engineering curricula.

- Utilities are developing their own A/I tools rather than relying exclusively on commercial products. Reasons for this are that existing tools are judged to be of little use in knowledge acquisition, that evaluating commercial products is time-consuming, and that many vendor products require a long learning curve [2]. Another factor is that the nuclear industry needs tools that combine symbolic and numerical processing. Functions for which the nuclear industries are developing tools include knowledge base construction, knowledge representation, the merging of numerical and symbolic processing, and the construction of plant models.

- Few expert systems are being developed for the express purpose of capturing human expertise. Perhaps this reflects the high level of training that all operators receive. As a result, no one individual stands out as an expert. Another consideration undoubtedly is that regulations require reactor operators to follow detailed, written procedures. Improvisation is not desired. Specific applications for which the capturing of human expertise is a prime factors motivating the nuclear industry to develop expert systems, (2) areas of application, and (3) issues related to the acceptance of expert systems within functioning power stations.

MOTIVATION FOR THE USE OF NUCLEAR EXPERT SYSTEMS

Expert systems are a special type of computer software for which the objective is to reproduce the capabilities of exceptionally talented humans. This is achieved by encoding human experience in various knowledge representation schemes. The nuclear and chemical industries have recently extended the concept to include reasoning about physical systems using information derived directly from the structure and function of those systems. The underlying idea is to design the expert system so that the experience of the human experts and the information on plant structure (the knowledge base) are kept separate from the method by which that experience and information is accessed (the inference engine). Expert systems differ from conventional algorithmic programming in two respects. First, as new information is obtained, it can be added to the knowledge base without revising the inference engine. That is, no reprogramming is needed. Second, an expert system can at any time provide the rationale for its conclusions. It does this by keeping track of the chain of deductions that support each particular conclusion.

The reasons for applying expert systems to the design, management, and operation of nuclear power plants are the same as for using them in business, medicine, or manufacturing. Namely, expert systems can assist in management, in diagnosis, and in the formulation of decisions given either uncertain or incomplete information. The emphasis here is on the word 'assist'. Expert systems, at least as presently constructed, are not a substitute for a human. They are, like any other tool, a means by which an already knowledgeable human can increase his or her productivity and efficiency.

Much of the appeal of expert systems to the nuclear industry originates with the structure of those systems. Expert systems are, as noted, very simple entities consisting of a knowledge base, an inference mechanism, and a user interface. For many nuclear applications, one must also add a component for the real-time acquisition of data. At its most basic level, an expert system is a means of performing automated searches. For example, the knowledge base may contain a set of production rules that are in the form 'if condition A and condition B are present, then the following regulation applies'. The function of the expert system is first to identify the current plant condition and then, via its inference mechanism, to compare the antecedent clauses of each production rule against the observed plant status. If a match exists, the rule is taken as applicable. The major advantage to this approach is that the knowledge base and the inference mechanism, which may be thought of as the software's main program, are separate. For the nuclear industry this means that as the plant's layout is changed or as new regulations are imposed, the knowledge base can be updated without incurring the need to revise the inference mechanism. Were a conventional programming technique to have been used, the entire program would require revision because the knowledge and the method for its interpretation would be intertwined.

Another feature of the expert systems approach that the nuclear industry finds appealing is the capability of the methodology to generate an explanation for its conclusions. Specifically, once a particular action has been identified as being appropriate, the system can print out a statement to the effect that such an action is required because the observed conditions exist. Moreover, it can cite the relevant supporting regulations. This feature is of particular use in the case of nested production rules where the presence of a certain condition may invoke a regulation that in turn makes applicable some other rule. Most regulatory codes are unfortunately written in such a manner and contain multiple interacting sub-

Table 1

APPLICATIONS OF EXPERT SYSTEMS WITHIN THE NUCLEAR INDUSTRY

Category	Number of Systems by Nation			
	France	Japan	U.S.	Other
Engineering Tools		4	15	2
Systems that Capture Human Expertise	2	4	3	2
Plant Design	2	5	11	3
Plant Management	5	8	13	3
Maintenance Applications	5	6	18	6
Interactive Diagnostic Systems		2	10	2
Real-Time Diagnostic Systems	6	12	20	5
Decision Support Systems	8	20	22	9
Emergency Preparedness and Response			13	3
Operator Behavior and Models	1	1	2	1
Control		8	15	2
Evaluations of Expert Systems		1	3	4
Totals	29	71	145	42

objective include training, the servicing of diesel generators, structural analysis, and the design of various plant components including electromagnetic pumps, manipulators, and heat exchangers.

- Several expert systems have been developed to assist engineers with the design of nuclear power plants and their associated interfaces to an electric power grid. Applications within this category have been quite varied and they constitute only a small fraction of the total. One of the more common applications of this type is for an expert system to assist in the execution of the large computer codes that are used for plant safety analysis. For example, the expert system might provide advice concerning both the modeling of the reactor core andthe interpretation of the code's output. Other applications include the design of electric distribution networks, the layout ofelectrical substations, pipe routing and support, and probabilistic risk assessment (PRA) studies. Relative to the last of these applications, it is noteworthy that there is an active exchange between PRA analysis and expert systems technology. Expert systems are used to assist in the construction of fault trees for PRA studies and the knowledge contained in existing fault trees is often used as the basis of an expert system.

- A number of expert systems have been developed to assist in the management of nuclear power stations. The objective here is to assure regulatory compliance. For example, the expert system could be used to match plant conditions against technical specifications and determine which were currently applicable. (Note: Technical specifications are a set of rules which define the plant operating conditions that must be maintained in order to ensure that the plant is at all times operated within the envelope of conditions analyzed in its Final Safety Analysis Report or FSAR. Technical specifications are part of a reactor's operating license and have the force of law.) Expert systems of this type need not operate in real time and their fields of search are known because the sets of regulations, although complex, are finite. Other managerial tasks for which expert systems are being developed include the generation of system tagouts and work authorizations, compliance with welding specifications and quality assurance standards, inspection programs including the identification of trends, plant life extension, the management of noise analysis codes, and rod pattern planning for boiling water reactors.

- Maintenance is another area for which a significant number of expert systems have been developed. Specific applications include spare parts inventory, the scheduling of repairs and calibrations, guidance on the servicing of valves and pumps, the planning of refuelings, steam generator inspections, the monitoring of radiation safety, and non-destructive testing. Maintenance expert systems, while similar to those for plant management, differ in that they often provide advice. For example, a system for the scheduling of repairs might provide an estimate of the remaining useful life of a component that is showing the incipient signs of wear.

- Interactive diagnostic systems are being developed for the analysis of physical processes that vary slowly. The challenge here is that the field of search may no longer be known. Applications include water treatment and cover gas analysis, the identification of the cause of plant trips, and the monitoring of plant thermal performance.

- Real-time diagnostic expert systems are currently at the cutting edge of the technology. Not only may the field of search be unknown, but there must be a direct data link between the plant and the system so that real-time analysis can be performed. Within this category are turbine generator diagnostic systems, such as GenAID, which have proven to be of significant economic value [3]. However, those successes notwithstanding, it is clear that the application of expert systems to diagnostics in general requires further research. For example, suppose that the system's knowledge base is inadequate and that as a result it can not achieve a correct diagnosis. Will that be obvious to the user? Or will the system provide an incorrect analysis that has all the appearances of being correct? In addition to turbine generator diagnostics, applications include loose parts detection, noise analysis, signal validation, alarm diagnosis and filtering, plant status monitoring, and causal analysis.

- Operator adviser and emergency response expert systems constitute about 25% of the total. These range from narrowly focused French systems for the operation of chemical and volume control systems to extremely broad Japanese systems intended for plantwide use [4-5]. In general, the more focused a system, the greater its likelihood of success. However, success can also be assured by careful design of the man-machine interface. This is the approach being taken by both Japan and Canada. The design of expert systems for decision support is a most challenging task because the systems must not only generate accurate analyses but they must also present those analyses in a manner that reinforces an operator's existing cognitive approach to plant operation. Otherwise, the operator will not use the system. Most decision support expert systems are for general diagnostics. However, there are specific applications in the areas of xenon oscillations, crane malfunctions, decay heat removal, procedure tracking, procedure generation and verification, and the operation of chemical and volume control systems.

- The rule-based approach and 'fuzzy' logic are being used by some researchers as a method for modeling operator behavior. Systems of this type constitute only a small fraction of the total being developed within the nuclear industry. The more important relation between expert systems and models of operator behavior is the incorporation of cognitive models in the expert systems. For example, this is being done as part of Japan's program 'Advanced Man-Machine System Development for Nuclear Power Plants' (MMS-NPP) [6]. The objective is to improve the man-machine interface.

- Research on the use of expert systems for reactor control is quite active, particularly in Japan and at certain universities such as the Massachusetts Institute of Technology. Rule-based control is seen as offering the possibility of robustness because the control action would be the net result of many rules, each linking the output of a particular sensor to a desired action. The combined effect of these rules renders the system insensitive to the loss of an individual sensor. The use of a rule-based system for the actual control of a research reactor has been demonstrated [7]. Moreover, it should be noted that many of the tasks being undertaken at the prototype level in Japan are those that will be needed for fully-automated, closed-loop control to be implemented on a plant-wide basis.

- Quantitative evaluations of the benefits of expert systems to reactor operators have been performed at both the Idaho National Engineering Laboratory (INEL) and at the Halden Project in Europe. The former involved assessing the benefits of an expert system as an operator aid during an emergency [8]. The latter was a comparison of expert and conventional alarm filtering systems [9]. Neither study showed any overwhelming benefit to the use of the expert system. The INEL study found that operators would not use an expert system to perform a task that they could accomplish directly by examination of plant instrumentation. The Halden study indicated, but did not conclusively demonstrate, that the expert approach to alarm filtering would be of benefit during major emergencies. Perhaps the only definitive conclusion that can be drawn about quantitative evaluations of expert systems is that there have been far too few of them.

Are the nuclear industry's expectations for the use of expert systems realistic? As yet there have been few actual implementations. What evidence there is suggests the presence of both positive and negative trends. As for the positive, some systems are in actual use. These include the French systems CERBERE and TIG which are for assistance with refueling and welding respectively. Also in France, the system EXPERT-GV is being used to train personnel in the identification of steam generator tube defects and the alarm filtering system EXTRA has been installed at a commercial site. Italy reports that the water chemistry monitoring system ERICE is functional. Certain components of the Japanese undertaking MMS-NPP are operational as are some of the systems for assisting reactor operators with the functioning of boiling water reactors. In the United States, systems for plant thermal performance monitoring, turbine generator diagnostics, and the generation of work permits have achieved commercial success. Also in the United States, the Alarm Filtering System (AFS) is in use at a fuel reprocessing facility. (Note: Details and reference information on these and related systems are given in [1].) The above list is by no means complete. Also, it can be expected to increase significantly over the next twelve to eighteen months as systems now completing prototype-testing become operational. Of significance is that the systems that either have achieved or are approaching commercial implementation cut across the spectrum of applications. Countering these positive developments are the experimental evaluations at both the Idaho National Engineering Laboratory and at the Halden Facility [8,9]. The results of those tests were at best inconclusive as regards the value of expert systems to reactors operators. Also, a most disturbing trend is that some of the systems that have completed prototype-testing have been shelved following brief in-plant trials. In summary, even if an expert system functions properly in a technical sense, commercial success is not assured.

ACCEPTANCE OF NUCLEAR EXPERT SYSTEMS

Why do some systems succeed while others fail? As originally conceived, the intent of an expert system was to make heuristic or experiential knowledge obtained from truly outstanding individuals available to everyone working in the field. Moreover, those systems were to be used in an interactive manner with the system querying the user for additional information. It is apparent that nuclear applications in the areas of plant design, plant management, maintenance, and interactive diagnostics generally conform to those criteria. However, applications in the areas of real-time diagnostics, decision support, emergency response, and control do not. The principal difference is that the latter require real-time solutions and entail the use of numerical models or other forms of 'deep knowledge'. These features are sometimes cited as being inappropriate for an expert system. It is true that their presence may make the construction of an expert system more difficult. However, they are certainly not the deciding factor in determining the likelihood of a system's ultimate success. In particular, there are numerous reports in the literature of prototype tests in which the real-time aspects of such systems have been successfully demonstrated. Moreover, some of the systems that either have achieved or are approaching commercial success are of this form. The practical extension of expert systems technology to real-time use and the incorporation of numerical models in those systems is something in which the nuclear (and also chemical) industries should take pride.

A better indicator of the factors that account for a system's acceptance and hence success can be obtained by examining the characteristics of those systems that are in commercial use. The sample base is admittedly small. However, it appears that commercially successful systems exhibit the following traits:

(1) The intended users of the expert system are generally not reactor operators. Rather, they are plant managers, welders, chemists, Q/A supervisors or startup engineers. This may be an advantage in that, unlike reactor operators, these user groups tend to be highly

defined. Hence, the design of the man-machine interface may be simpler.

(2) The systems being developed are for the purpose of assisting, not replacing or supplanting, a human. The objective is to improve productivity by giving the user more immediate access to necessary information.

(3) Many areas of application are highly focused. This limits the extent of the knowledge base needed to support the system. That in turn means that many issues related to the system's construction and implementation are simplified.

(4) If the area of application is broad, then substantial emphasis is placed on the quality of both the knowledge base and the man-machine interface. This is true of both the turbine generator diagnostic systems and of many of the Japanese systems.

(5) Regulatory issues are less of a concern because a human remains in overall control and makes the final decision.

Assuming no technical deficiencies, the issue most crucial to the acceptance and hence commercial success of a nuclear expert system appears to be the man-machine interface. This involves much more than a well-conceived graphics display although that too is of importance. The question is whether or not the system truly supports the user. In particular, does the expert system provide the user with the information that he or she needs? Does it do so in a manner that reinforces the operator's existing cognitive processes? Or is the operator forced to alter his or her pattern of thought in order to conform to the system's mode of deduction? Does the knowledge base reflect the true complexity of the plant? Or must the operator make allowances for limitations in the expert system's advice? Is data acquisition automatic? Or must the operator supply information to the system? These are the fundamental questions that govern a system's acceptance and use. Another issue of importance is that of regulatory acceptance.

Listed below are some of the factors relevant to the acceptance of a nuclear expert system:

- The system should provide the user with the information that he or she needs. Moreover, extraneous material should not be forced on the user. Relative to licensed reactor operators, the need is for real-time, accurate diagnostics. Operators are highly trained professionals and it would be most unusual for an operator not to be aware of the appropriate action once plant status is known. For example, the problem at Three Mile Island was that the operators did not recognize the plant's true condition.

- Expert systems systems should be designed to support an operator's cognitive processes and to reinforce the operator's existing approach to plant operation. For example, experienced operators use pattern recognition skills to monitor plant behavior. Yet, many expert systems use a deductive mode of reasoning. Does it make sense to require the operator to conform to the machine's method of analysis?

- The limitations associated with an expert system should be obvious. Otherwise, the user will have to supervise the machine. Moreover, the operator will be placed in the difficult position of having to decide between his or her own judgment and the machine-generated advice.

- If an expert system is to be used by several different groups (e.g., reactor operators, senior operators, shift technical advisers) then multiple interfaces should be designed. Each interface should reflect the expectations, education, and skill levels of its assigned user group.

- Displays should be uncluttered and use easy-to-read, high quality graphics.

- Real-time adviser expert systems should exhibit the same relation to an operator as do reactor instruments. That is, the requisite information should be continuously displayed and the operator need only look at the display screen to obtain an update.

- Expert systems intended for diagnosis and operator support should not involve the operator in the process of data acquisition. Rather, the expert system should obtain the requisite information from the plant process computer and/or directly from the sensors.

There are of course many other factors involved in the acceptance and success of nuclear expert systems. These include the content and organization of the knowledge base, the ease with which the system can be updated, the presence of the instrumentation needed to provide raw data, the computer aptitude of the prospective user, the problem chosen for solution by the expert system, and regulatory attitudes. These and other factors are discussed in detail in both the book [1] and in a related review [10].

SUMMARY AND CONCLUSIONS

In summary, expert systems technology has the potential to make a significant contribution to the reliable operation of nuclear power stations. Moreover, that potential will probably be realized in certain areas related to plant management such as compliance with regulations and the performance of diagnostic tasks that can be done interactively. However, it remains an open question as to whether expert systems can be successfully applied to other areas including real-time diagnosis and guidance. For this to happen small-scale demonstrations that clearly illustrate the utility of the technology must be performed. Also, many issues related to the effective design of the man-machine interface must be identified and resolved. This is an enormous challenge because, despite much excellent research on the topic, there is undoubtedly much that we still do not know. Also, in the final analysis, the only acceptable means of verifying system effectiveness will be through actual testing under as realistic conditions as possible. In the interim, both the nuclear and the A/I communities should resist the urge for immediate implementation and instead adopt an incremental approach whereby steady progress is made towards rendering the technology truly effective.

ACKNOWLEDGMENTS

Appreciation is expressed to Ms. Georgia Woodsworth and to Mr. Ara Sanentz for their assistance with manuscript preparation. This paper is based on a book published by the authors through the American Nuclear Society [1].

REFERENCES

1. J. A. Bernard and T. Washio, Expert Systems Applications Within the Nuclear Industry, American Nuclear Society, LaGrange Park, IL, July 1989.

2. D. Hu, "Expert System Applications in Energy Management and Utilization at EPRI," Proc. EPRI Seminar on Expert Systems Applications in Power Plants, Palo Alto, CA, Electric Power Research Institute, May 1987.

3. R. L. Osborne and M. Coffman, "Power Plant Experience with Artificial Intelligence Based, On-line Diagnostic Systems," Proc. Am. Nucl. Soc. Int. Topical Meet. on Artificial Intelligence and Other Innovative Computer Applications in the Nuclear Industry, D. Majumdar and J. Sackett, Eds., Plenum Press, New York, NY, pp. 547-553, 1988.

4. J. Guillermard and J. C. Lagache, "PMS-1: A Real-Time Advisor for the Boron Control in PWR Nuclear Reactors," Robotics (Netherlands), Vol. 22, No. 4, pp. 323-328, 1986.

5. T. Ogino, Y. Nishizawa, T. Morioka, N. Naito, M. Tani, and Y. Fujita, "Intelligent Decision Support Systems for Nuclear Power Plants in Japan," Reliab. Eng. Syst. Saf. (U.K.), Vol. 22, No. 1-4, pp. 387-399, 1988.

6. Y. Fujita, "MAPI's Knowledge-Based System Will Help to Deal with Abnormal Conditions," Nucl. Eng. Int. (U.K.), Vol. 32, No. 396, pp. 28-31, 1987.

7. J. A. Bernard, "Use of a Rule-Based System for Process Control," IEEE Control Systems Magazine, Vol. 8, No. 5, pp. 3-13, 1988.

8. M. A. Bray, "Evaluation of an Artificial-Intelligence-Based Operator Aid," Trans. Am. Nucl. Soc., Vol. 46, pp. 293-294, 1984.

9. E. C. Marshall, C. S. Reiersen, and F. Owre, "Operator Performance with the HALO II Advanced Alarm System for Nuclear Power Plants: A Comparative Study," Proc. Am. Nucl. Soc. Int. Topical Meet. on Artificial Intelligence and Other Innovative Computer Applications in the Nuclear Industry, D. Majumdar and J. Sackett, Eds., Plenum Press, New York, NY, pp. 353-360, 1988.

10. D. Majumdar (Ed.), Artificial Intelligence Applications in the Nuclear Industry, DOE/ID-10191, Oct. 1988.

Water Chemistry Expert Monitoring System

A. J. HARHAY and N. C. LEONI
Rochester Gas and Electric Corporation
Rochester, New York, USA

S. G. SAWOCHKA and S. S. CHOI
NWT Corporation
San Jose, California, USA

ABSTRACT

Rochester Gas & Electric Corporation has initiated demonstration of an artificial intelligence (AI) expert system for the on-line monitoring and diagnosis of secondary water chemistry at the Ginna Nuclear Plant. The Water Chemistry Expert Monitoring System (WCEMS) is a PC based expert system integrating data acquisition, chemistry analysis, and expert system software. Using the output from 26 in-line sensors, WCEMS continuously reviews the water quality to augment the conventional chemistry monitoring program. Maintaining the excellence of secondary water chemistry control is critical to minimizing the potential for steam generator corrosion problems. The rapid identification of impurity ingress and initiation of corrective actions are essential to insuring safe operation and maintaining the long-term integrity of secondary system components.

INTRODUCTION

Rochester Gas & Electric Corporation (RG&E) has recently initiated demonstration of an internally funded research & development project that applies artificial intelligence (AI) technology in developing an on-line expert system for continuously reviewing and diagnosing secondary-side water chemistry conditions at the Ginna Nuclear Power Plant (1, 2). This application involves the acquisition of real-time data from 26 in-line instruments used to characterize feedwater, steam generator, and steam circuit chemistry conditions. The WCMES consists of three networked PC subsystems, data acquisition, data analysis, and expert subsystems. The maintenance of stringent chemistry controls and the early recognition of potentially detrimental conditions are critical to minimizing the corrosion of tubes in Ginna's steam generators. The WCEMS application was pursued for the benefits that could be provided in overall chemistry control and also because it was felt that this relatively small application could serve as an effective forerunner project for gaining experience with the technology.

RG&E is working with the NWT Corporation (San Jose, CA) and the Electric Power Research Institute (EPRI, Palo Alto, CA) in the development of the WCEMS application. NWT is the principal contractor, bringing to the project expertise in both power plant chemistry and computerized data assessment techniques. They have provided the hardware, software, and extensive support in structuring the application. The expert system software is the EPRI-developed Small Artificial Reasoning Tool (SMART), an AI software for PCs which was designed not simply as a "shell", but as a "toolkit" for building an expert system (3). EPRI is providing user programming support and the necessary technical support for effectively integrating SMART into the system.

Presently, the WCEMS project is entering a second stage of field testing after the implementation of enhancements identified during testing in the fall of 1988.

PLANT DESCRIPTION

The R. E. Ginna Nuclear Plant is a single pressurized water reactor unit with a Westinghouse nuclear steam supply system which has two coolant loops and two recirculating steam generators. The plant began commercial operation in March 1970. A secondary water circuit schematic is shown in Figure 1.

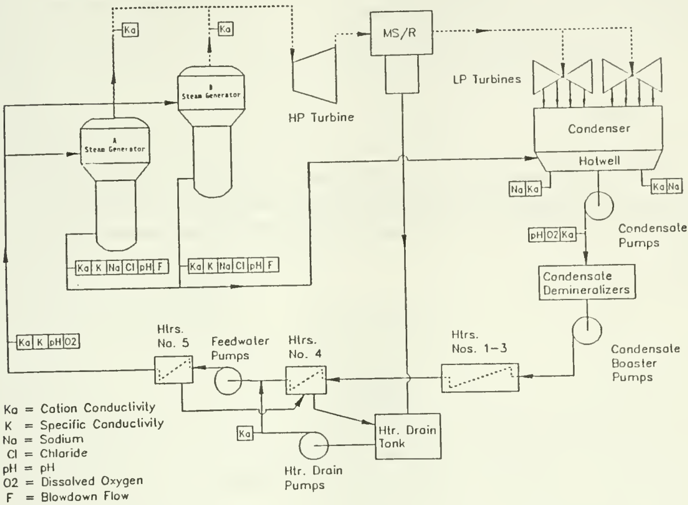


Figure 1 Ginna Secondary Water Circuit

Steam from the two steam generators is expanded through the high pressure (HP) turbine from which it exhausts into moisture separator/reheaters. Reheated steam is passed through two low pressure (LP) turbines. The condensate pumps which take suction from the condenser hotwells discharge to a deep-bed condensate polisher system. Polished condensate flows through several coolers/condensers and then through two parallel strings of low pressure and high pressure feedwater heaters.

Both in-line instrument monitor readings and grab sample analyses are employed to characterize secondary water chemistry. The type and sample location of the in-line monitors used by the WCEMS are shown in Figure 1. Continuous measurements of cation conductivity, specific conductivity, sodium, chloride, pH, dissolved oxygen, and blowdown flow from various locations are centrally available as meter and strip chart displays at the secondary chemistry panel in the turbine building. These measurements readings exist as meter and strip chart displays. Data acquisition from the polisher influent, polisher effluent, individual polisher beds and makeup demineralizer plant was not pursued in the present project although the WCEMS is capable of handling such inputs.

WCEMS DESCRIPTION

The installed system consists of A/D convertor & transmitter hardware and three PCs for performing data acquisition, data analysis, and diagnostic reasoning. The configuration of the system installed is shown schematically in Figure 2.

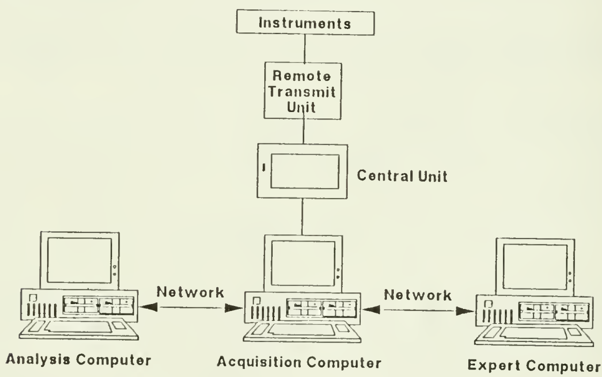


Figure 2 Water Chemistry Expert Monitoring System

The WCEMS was modularly designed so that the application for acquisition, analysis, and diagnosis could be built and operated independently. The integration of the System was developed using a file transfer of communication, as opposed to program-to-program data transfer. The potential benefits of upgrading the computer hardware are being considered. The three PCs are networked via IBM PC network hardware and Novell Netware software. The data acquisition computer also functions as a nondedicated network file server.

DATA ACQUISITION SUBSYSTEM

The data acquisition subsystem is comprised of the following components:

- Molytek 32-channel Remote Transmit Unit (NEMA4)
- Molytek 2702-C Central Unit
- Compaq Deskpro 386 Personal Computer
  (6 MB RAM, dual 5-1/4" flexible disk drive, and 60 MB fixed disk)
- Sony Color Monitor (high resolution graphics)
- IBM PC Network Adaptor
- Novell Netware software
- DOS operating software
- Molytek Molygraphics data acquisition software

The analog output signals of the in-line monitors are directly connected to the remote transmit unit (RTU), located near the chemistry panel in the turbine building. The RTU sequentially polls each instrument and converts the analog signal into engineering units to build a data scan set from all 26 monitors. The RTU may be programmed from the central unit, that is, the scan set is defined by assigning each monitor a channel number, a tag or label, an algorithm for conversion to engineering units, a unit of measurement, and an alarm set point.

Upon completion of signal conversion, each data scan set is transmitted to the central unit located in the secondary chemistry laboratory via an asynchronous RS-232 interface. The central unit displays the time of day, input values with units, and alarm status of each channel on a 32 character digital display. The central unit may also print a data log and/or trend plot on chart paper. Trend plots of any input channel parameter can be selected while the unit is in operation. The central unit transmits the data scan set to the acquisition computer via an RS-232 interface.

The data acquisition computer is located in the secondary chemistry laboratory near the central unit. Molygraphics (MG) software receives scan sets from the central unit about every 2 seconds. Scan sets are stored at a user defined frequency (at Ginna once every 6 minutes) and builds MGDATA files. An instantaneous data file (SIMOFILE) is update at a user defined frequency and can be accessed by the expert system for the diagnosis of secondary chemistry. Tables, trend plots, and bar charts of scan sets can be displayed using MG software. Another feature of MG is the "run back" which allows scan sets to be saved at a faster frequency for a certain time interval prior to and after an alarm occurrence. The user views the various MG display, plots, or charts via user developed menus. The user may flag or tag out a monitor during calibration, maintenance, or periods of malfunction, so that the expert system does not utilize the data in the review process.

DATA ANALYSIS SUBSYSTEM

The data analysis subsystem is comprised of the following components:

- Leading Edge Model D Computer
  (640 KB, RAM, 5-1/4" flexible disk drive, and 30 MB fixed drive)
- Sony Color Monitor (high resolution graphics)
- HP Ink Jet Printer
- HP 6-pen Graphics Plotter
- IBM PC Network Adaptor
- Novell Netware software
- DOS operating software
- NWT Data Analysis software

The data scan sets collected and stored by the acquisition subsystem in MGDATA files are transferred to NWT data files by using copy subroutines included in the NWT data analysis software and Molytek's conversion program MG123. The transfer to NWT data files provides data reduction and integration with manual entered data. The data reduction is accomplished by stripping out unused channels, any 'tagged out' monitors, and scan set header information. The NWT data files are utilized as the working data base and the MGDATA files as the archival data base.

The NWT data analysis software provides the capability to manipulate all stored data (both on-line and manual entry) and can present the results in several different graphical and tabular formats. Drawing upon the data base, short term and long term trends may be displayed on the screen or sent to a plotter. Tabulated data summaries can be displayed on the screen, as well as, output to a printer. Manipulation of individual variables or combinations of variables is possible for verification of data consistency and assistance in correlations. Summary histograms can be developed from the stored data to clarify variations in system chemistry and provide statistical analyses, i.e., average, minimum and maximum values, standard deviation, etc.

EXPERT SUBSYSTEM

The expert subsystem is comprised of the following components:

- Leading Edge Model D Computer
  (640 KB RAM, 5-1/4" flexible disk drive, and 30 MB fixed drive)
- IBM color monitor (low resolution graphics)
- IBM PC Network Adaptor
- Novell Netware software
- DOS operating software
- MULISP software
- SMART software

The expert subsystem receives a data scan set from the acquisition subsystem via the SIMOFILE and emulates the reasoning processes of a knowledgeable chemist to identify and diagnose abnormal chemistry conditions and provide advice, i.e., corrective action steps. The structure of the expert system is shown in Figure 3.

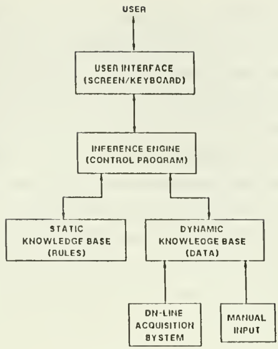


Figure 3

SMART was used as a tool to build the expert system. The SMART software is intended to serve as a primer for expert system concepts and to provide an environment that supports modest applications. It was selected as the AI software because of code capabilities relative to RG&E short and long range goals and EPRI's willingness to provide technical support for implementation. It should be noted that SMART has been developed from KEE, a much larger expert system program developed for industry use. Applications of this program are already being pursued by several utilities, which should facilitate utility interfaces for addressing other RG&E areas of possible AI application. The software provides for:

- Frame based knowledge representation with inheritance properties
- Forward and backward chaining reference methods
- Embedded functions
- Query functions
- Explanation capabilities

The WCEMS's knowledge base, i.e., data base, consists of both static and dynamic data, as shown in Figure 3. The static knowledge base contains the chemist's reasoning logic used to identify problem conditions. This knowledge base is developed in the English-like symbolic language of LISP in the form of "rules" which are easily understood by non-computer specialists. For example, the rules developed to establish the presence of a condenser leak are given in Figure 4.

PROBLEM: CONDENSER LEAK

- IF SODIUM-HWA RATE-OF-CHANGE IS > LIMIT OF 0.5 PPB/hr
- IF CATION CONDUCTIVITY-HWA RATE-OF-CHANGE IS > LIMIT OF 0.005 umhos/cm/hr
- THEN CONDENSER "A" LEAK IS CONFIRMED

Figure 4

The dynamic knowledge base contains the data scan set values, calculated rate of change and running average rate of change, identified conditions (if the identified conditions have been acknowledged), and date and time of last scan set read.

Two approaches are presently employed to evaluate secondary water chemistry at Ginna. First, absolute values of key parameters are continuously compared to action level values and the limiting specifications. Action levels and their associated chemistry limits were developed by the industry to define minimum requirements for system protection. A total of 46 rules were employed for the absolute value diagnosis. The limiting secondary chemistry specifications used in the knowledge base are given in Table 1.

Table 1

LIMITING SECONDARY CHEMISTRY SPECIFICATIONS*

PARAMETER	ACTION LEVEL 1			ACTION LEVEL 2			ACTION LEVEL 3		ACTION LEVEL 4
	CONDENSATE	FEEDWATER	BLOWDOWN	CONDENSATE	FEEDWATER	BLOWDOWN	CONDENSATE	BLOWDOWN	BLOWDOWN
pH @ 25°C	<8.8 or >9.2	<8.8 or >9.2	<8.5 or >9.0	N/A	<8.8 or >9.2	<8.5 or >9.0	N/A	N/A	N/A
Cation Conductivity uMHOS/cm @ 25°C	>0.15 but <0.3	>0.1 but <0.2	≥0.20 but <0.8	N/A	≥ 0.2	≥0.8 but <2	N/A	>2 but <7	≥ 7
Sodium, PPB	N/A	N/A	>3 but <20	N/A	N/A	≥20 but <100	N/A	≥100 but <500	≥500
Chloride, PPB	N/A	N/A	>5 but <20	N/A	N/A	≥20 but <100	N/A	≥100 but <500	≥500
Dissolved Oxygen, PPB	>10	> 5	N/A	>10 but <50	> 5	N/A	> 30	N/A	N/A

---

* R.E. Ginna Secondary Water Chemistry Monitoring Procedure No. WC-15

A second set of diagnostic rules was constructed based upon the average rate of change of impurity conditions, e.g. steam generator chloride => 0.5 PPB/hr, condensate pH=>0.05 UNITS/hr, etc. A third series of rules is presently being developed relating to the response consistency of monitors. A series of scenarios were developed for the most common problem conditions which could be identified by rate of change values. Currently, eight specific problem cases can be evaluated, utilizing the static and dynamic knowledge bases. Additional problem conditions are to be added in the near future.

The expert system execution cycle is as follows:

1. Read a scan set into the data dictionary from a copy of the SIMOFILE.
2. Convert the data dictionary ASCII string values to numeric values.
3. Calculate the rate of change and running average rate of change.
4. Run the backward chainer.
5. Display any identified problem conditions on the screen and store them in a event log file.
6. Accept a user interrupt to acknowledge the conditions and store the acknowledgement in the event log.
7. Display corrective action steps.

The system is currently being refined to make the advisory feature, i.e., the corrective action steps more user-friendly. The advisor would correlate actions with each individual problem case and would organize the actions on a priority basis. An example of an advisory for a parameter exceeding Action Level 4 is shown in Figure 5.

DATE: 89, 03, 09  TIME: 14, 19, 40

- ADVISOR -

1. IMMEDIATELY VERIFY THAT MONITORS ARE FUNCTIONING PROPERLY.
2. IF FUNCTIONING PROPERLY, INFORM CHEMISTRY & OPERATION S SUPERVISION OF ACTION LEVEL CONDITION.
3. REQUEST MAXIMUM BLOWDOWN FLOWRATES.
4. VERIFY READINGS WITH LAB METER & INFORM CHEMISTRY & OPERATIONS SUPERVISION.
5. PER WC-15 SPECS, CONFIRMATION OF ACTION LEVEL 4 REQUIRES SHUTDOWN WITHIN 4 HOURS. CHEMISTRY SUPERVISION WILL ADVISE TO THE APPROPRIATE CLEANUP MEANS.

Figure 5

The system also is being developed to provide a training tool aimed at enhancing the ability of technicians to understand and deal with chemistry transients. For training, simulated chemistry conditions would be entered into the dynamic knowledge base by using the keyboard. Technicians would predict specific problems for each simulated chemistry condition and compare their results with the results given by the expert system. Also, the training tool will hopefully provide a means of verifying and validating the expert system prior to final acceptance.

SYSTEM COST AND BENEFITS

The WCEMS is RG&E's and NWT's first venture into AI expert system development and, partly for that reason, a major portion of the funding is being provided by the RG&E Research and Development Committee. The total cost of the project will be approximately $160,000. This includes the hardware and software associated with each subsystem, RG&E and NWT labor for developing the application and structuring SMART, and plant modifications made to provide conductivity outputs that would properly interface with the acquisition system. This project also represents EPRI's first use of SMART in an on-line mode.

For RG&E, an important spin-off from the project will be the knowledge gained by their people in expert system development--knowledge which can be applied to future AI projects supporting other operations in the company. As a first-of-a-kind effort for RG&E, the project is expected to attract considerable attention and hopefully stimulate ideas for other applications. Although gaining experience in expert system development is an important goal, the first objective of the project is to further improve secondary water chemistry control at Ginna.

Almost all pressurized water reactor plants have experienced tube corrosion in their steam generators. Of the 23 U.S. steam generators similar to Ginna, 15 have already been replaced or extensively repaired. This is a enormous undertaking, with associated costs generally over $100 million per plant. Ginna is also experiencing tube corrosion, but fortunately the rate has been low enough that replacement has not been required. While careful attention to maintaining water chemistry control in the past is believed to be a significant factor in limiting tube corrosion at Ginna, it is recognized that even more stringent controls and faster response to off-normal water chemistry conditions will likely be required to minimize future problems.

The primary benefit of the WCEMS to RG&E will be in its potential to provide an overall improvement in chemistry monitoring, data interpretation and response to developing conditions. Until implementation of WCEMS, the recognition of hour-to-hour and day-to-day trends in chemistry parameters depended on a chemist or technician periodically reviewing the data on a strip chart recorder in the plant. Depending on a variety of factors, such as chart speed and the number of points being tracked on a single chart, the ability to note subtle trends can range from difficult to very difficult; and, of course, the retrieval of past data from charts is a tedious challenge. With the incorporation of matrices utilizing rate of change criteria, as well as warnings at various absolute values, the expert system can reason that something is happening and provide advice to the technicians and operators in a time probably faster than "humanly" possible. Prompt action to minimize the extent of a chemistry transient can potentially minimize tube degradation, thereby reducing the extent of subsequent repairs and prolonging the useful life of the steam generators.

Use of the WCEMS for on line data review also will strengthen the Ginna chemistry program by providing a cost effective, round-the-clock diagnosis of chemistry conditions by capitalizing on the expertise of senior chemistry personnel. Hiring experienced chemists to enable providing continuous expert review of chemistry data would likely cost about $200,000 annually...significantly more than the development cost for the WCEMS. In fact, with the WCEMS, RG&E hopes to be able to "save money" by somewhat freeing its human experts to acquire new knowledge and pursue new avenues for improving the quality of existing programs.

FUTURE DIRECTION

Assuming successful demonstration of the WCEMS, additional on line chemistry inputs will likely be added, e.g., makeup demineralizer and condensate polisher plant data. The networking of additional PCs also is envisioned to allow access to the acquisition system from other locations, such as the plant chemist's office, the plant auxiliary operator's office, and corporate chemistry offices in Rochester.

It also is anticipated that RG&E will pursue development of an on line expert system for use by chemistry and operational personnel at their fossil plants, as well as, investigate possibilities for applying AI expert system technology to other Company operations.

ACKNOWLEDGEMENTS

The cooperation of EPRI and Molytek, Inc. in the development of the Water Chemistry Expert Monitoring System is greatfully acknowledged. The extensive support and assistance of Dr. David Cain (EPRI) and Mr. Duane Ellis (Molytek) contributed immensely to the success of the project.

REFERENCES

1. Harhay, A.J., Choi, S.S., Cain, D.G., Leoni, N.C., Sawochka S.G. "Water Chemistry Expert Monitoring System", Discovery '88 Computer Forum, Engineering & Operating Division, Edison Electric Institute, Sept. 1988.

2. Leoni, N.C. and Harhay A.J., "Data Acquisition System Minimizes Corrosion in Nuclear Power Station", I&CS, Jan. 1989, p. 99-101.

3. Cain, D.G., Small Artificial Reasoning Toolkit, Electric Power Research Institute, Inc. 1988.

The Utility Experience of Implementing the Emergency Operating Procedure Tracking System

W. C. CHANG and J. F. CHENG
Taiwan Power Company
Kuosheng Nuclear Power Station
60 Pa-Tou, Yeh-Liu Village
Wan-Li Hsiang, Taipei Hsien, 20703
Taiwan, ROC

ABSTRACT

This report presents the experience of a project sponsored by the Electric Power Research Institute (EPRI),Taiwan Power Company (TPC)and supported by the Nuclear Software Service (NSS), General Electric Company (GE) and Science Applications International Corporation (SAIC) to implement the Emergency Operating Procedure Tracking System (EOPTS) in Kuosheng Nuclear Power Station Simulator. Before implement the EOPTS in Kuosheng simulator, the Safety Parameter Display System (SPDS) of the Emergency Response Facility Technical Data System (ERFTDS) shall be stimulated, the hardware and software linkage between the simulator and ERFTDS shall be established, that include installation of a VAX-8200 computer, Gould - Vax computer hardware linkage, ERFTDS software installation, simulator source variables selection and linkage it to the ERFTDS database.

SECTION 1

BACKGROUND

Over the past several years, the EPRI has sponsored projects in the area of "advanced operator aids" computerized system known as the IMAGE system. One of the applications of IMAGE system, the Boiling Water Reactor Advanced Operator Aids (BWR-ADA) version, is designed to use the plant parameters database obtained from the Hatch Simulator. But it is still too slow to be used in the online system. Over the last seven or eight years a significant efforts have been extended by the BWR Owners Group to develop the generic Emergency Procedure Gukdelines which are transfered into the plant specific Emergency Operating Procedures (EOPs). This project is to develop a more advanced and complete system using the high speed "C" language to perform the EOPTS in conjunction with the SPDS.

1.1 EMERGENCY OPERATING PROCEDURE TRACKING SYSTEM

The emergency operating procedure tracking system (EOPTS) is based on the emergency procedure guidelines (EPGs) revision 3L of the BWR Owners Group, using the Taiwan Power Company's Kuosheng Nuclear Power Station emergency operating procedures (EOPs) as a specific model. The system traveres the entire EOPs logic at short time intervals and provides an online display of the appropriate steps in these EOPs. By enhancing the operator's abilities to interpret and apply these procedures, the computer-based EOPTS developed by the EPRI can help to reduce the human error.

1.2 EMERGENCY RESPONSE FACILITY TECHNICAL DATA SYSTEM

The installation of the Emergency Response Facility Technical Data System is one of the requirements of U.S. NRC NUREG-0737, which provides online monitoring of the plant measured points (digital, analog and pulse) representing significant plant process variables. The system scans digital and analog inputs at a specified intervals, processes the data and provide various on-line display (such as safety parameters display), plots of current, predicted or historical plant performance and on-line/off-line logs of plant parameters.

The Safety Parameter Display System (SPDS) is one of the functions of the ERFTDS which provide a concise display of critical plant variables to the control room operators to aid them in rapidly and reliably determining the safety status of the plant. The principle purpose and function of the SPDS is to aid the control room personnel during abnormal and emergency conditions in assessing whether abnormal conditions warrant corrective action by operators to avoid a degraded core.

1.3 COMPANIES PARTICIPATE IN THE PROJECT

The companies participate in the project are as follows:

a. Electric Power Research Institute (EPRI), manager of the EOPTS development in U.S.A. and provide the EOPTs protocol.

b. Taiwan Power Company (TPC), handling the overall project in the Kuosheng simulator and final setup.

c. Nuclear Software Services (NSS), provide the EOPTS kernel program.

d. General Electric Company (GE), vendor of the ERFTDS, provide the Gould-Vax computer software linkage and EOPs rule logic.

e. Science Applications International Corporation (SAIC), provide the Gould-Vax computer hardware linkage.

f. Accident Prevention Group (APG), Coordinate the human cognitive reliability test.

1.4 OBJECTIVES OF THE PROJECT

The objectives of the project are as follows:

a. Develop the computer capability for the EOPTS.

b. Check and modify the EOPs rule logic and SPDS algorithms as nccessary to support the EOPTS.

c. Verify and validate the EOPTS for in-plant use at the Kuosheng Nuclear Power Station.

d. Prepare for the evaluation of EOPS by control room operators.

e. Transfer the experience and technology to the other utilities.

SECTION 2

EXPERIENCE OF IMPLEMENTING THE EOPTS

2.1 GOULD-VAX COMPUTER LINKAGE

The linkage was installed by the SAIC at March, 1987. The hardware linkage include HSD Card, HSD Cable Interface Card and DEC Compatible DMA Interface Card installation. The software linkage include the following steps:

a. Create a new SYSGEN directive file, this is normally accomplished by running the EDITOR, reading the existing SYSGEN directive file, inserting new lines to include the Q-LINK driver in the executive and then writing a new SYSGEN directive file.

b. Create a new COMPRESS input file, this is normally accomplished by editing the existing file.

c. Run LIBED to insert QSET into the MPXLIB.

d. Run COMPRESS to create the new object file for SYSGEN.

e. Run SYSGEN to creat the new executive.

f. Test the new executive and software linkage, the test program should be run on both the Gould and Vax machines.

g. Once the new executive test is finished, it will establish the bootable system on the system disk.

The simulator is failed to run after a user device U360 is assigned to the SYSGEN file. The driver OH.HSD30 was restored to the disk from the original HSD handler object tape, rerun the COMPRESS AND SYSGEN then the simulator was back to normal operation.

When performing the new executive and software link test, no communication between the Gould and Vax computers due to the test program provided by the SAIC has a mismatch revision. After the program in the Gould computer was modified, the test is satisfactory, the linkage speed is about 30,000 byte per second.

2.2 ERFTDS SOFTWARE INSTALLATION

The ERFTDS software was installed at April, 1987. The major job is to test the interface software between the Gould and Vax computers. The interface software provides an effectual method for transmitting the simulator data and status (Freeze, Run, Reset ... etc.) information to the Vax in place of the ERFTDS Data Acquisition System (DAS).

The interface software is composed of both online and offline functions. The online function gather the process data and status from the Gould computer simulator global memory and covert it into a formate that is compatible with the Vax computer, then transmit it to the Vax. The online function also receive the information from the Vax and respond back to the Gould appropriately. The offline function provide a method of generating and modifying the DAS signals without modification the source program of simulator, GEPAC plus or the interface software itself. A series of four (4) program generate mapping files are loaded by the online function during system initialization. The mapping files contain the information necessary to generate the data point buffer from the process data available in the simulator global memory.

The first step in preparation to run the Emergency Response Information System Sampler (ERISSAMP) is to generate the ERFTDS point configuration mapping files, A list of the ERFTDS points to be simulated must be established, the analog and digital point files (ER:APF and ER:DPF) are constructed from this list. The simulator source point for each ERFTDS point must be determined, the sampled analog and digital source files (ER:SASRC and ER:SDSRC) must be constructed, points that are not simulated must be specified as constant points then entered in the constant analog and digital source files (ER:CASRC and ER:CDSRC), these files shall be " stored " as system files. Each of the mapping program is then running to generate the mapping files.

The problems experienced during this phase were as follows:

a. The original offline program was based on the Datapool concept, but Kuosheng simulator software was based on the Simulator Software Support (S3) system developed by the Singer Link. The date base concept are quite different.

   The Datapool is a memory partition defined either at SYSGEN or via the File Manager utility (FILEMGR), it is structured via the datapool dictionaries that were built and maintained by the Datapool Editor (DPEDIT) which provides the ability to add, change, delet and equate variables in an existing dictionary or build a new dictionary. If a variable is changed, it will change the dictionary and all tasks which reference to the partition are simply recataloged with the modified dictionary.

   The S3 system supports the creation and usage of a sophisticated data base structure. It will satisfy a wide range of real time simulation applications and can be easily implemented on most computer system configurations. All simulator date, both variables and constants, are located in a common memory area accessable by all the simulation programs. The structure of the common memory area is created by using the global common machanism available in all standard FORTRAN compilers. The content and structure of the data base are defined by a Master Data Dictionary (MDD), which is created and modified under the control of Data Base Manager (DBM) program.

The Kuosheng software engineer had to developed a routine to open and read the MDD file data and modified all Datapool related off-line programs to enable access the MDD file to get the right data.

b. The interface program shall get the simulator status (Freeze, Run, Reset, ... etc.) and transmit it to the Vax DAS program for handling the condition, but the original interface program "SMSTAT.F" could not get the right status information, since it was based on another simulator software, so that it had to be modified.

c. The point composer is used to generate ERFTDS points, for which it did not have a corresponding simulator source point readily available, by programming an equation which may use numerous simulator variables. The program is entered as composition instruction similar to the assembly language and had to be modified for the GLOBAL memory usage since it is different from the Datapool concept.

2.3 SIMULATOR SOURCE VARIABLES SELECTION

The ERFTDS data points (about 2,000) were selected from the simulator database (about 20,000 points). The definition and engineering unit (analog points) or zero/nonzero status (digital points) of the ERFTDS data points were carefully studied, then select the corresponding variable name in the simulator database. If the ERFTDS data point were not simulated then the new point(s) were added and the associated simulator model should be modified to provide the dynamic input signal(s) to the ERFTDS. After the data points selection, the dynamic response were checked by running the simulator with the necessary operation condition set up.

2.4 EOPTS SOFTWARE INSTALLATION

The EOPTS software program was installed at March, 1988. The integration test of the EOPTS software is intend to verify the interface between the NSS software and the GE GEPAC+ system. It includes the ability to get information out of and into the Habitat point definition data base, the ability to start and stop the EOPTS, the ability to display EOPs massage on a dedicated VT220 terminal and change the color of EOP status box(es) on the SPDS monitor when the EOP entry condition(s) are meet.

The EOPTS failed to initial start-up after installation, that forced the Kuosheng software engineer to study the "C" language, data structure and kernel program then debug the whole system and modified the command procedure to set a correct data directory.

With plant simulator in normal operation, starting the EOPTS and runing the EOPs message clear function, the screen of the dedicated VT220 terminal shall display "NO MESSAGE" only, but it was fill up with lots of message. The LCPTGET subroutine for handling the dynamic data and the logic to get the process constant in the SETDATA.C program were incorrect and the "NOT" logic in the LOGIC.C program was incorrect too. The Kuosheng software engineers modified the SETDATA.C program to prevent it from tagging the dynamic data as a "BAD" data, to check if it is a process constant then skip to get the data in every cycle time, also debug the "NOT" logic in the LOGIC.C program to solved the above problems.

2.5 EOPTS DATABASE DICTIONARY VERIFICATION

The EOPTS database dictionary is maintained on the Vax computer as an ASCII file and contains the definition and value for each data point used in the EOP tracking system. The dictionary forms the linkage between the parameters used in the rules and the input parameters from the GE database.

The database dictionary includes the followings:

1. The parameters used in writing the rules, these parameters are points obtained from the GE database, variables derived within the rules and EOP logic states.
2. The corresponding name in the GE database, if it is an input parameter.
3. The data type of the parameter of variable.
4. The priority, if the parameter is an EOP state.
5. The address where the value is stored.
6. The message, if the variable is a state.
7. Quality tag.

The database dictionary received from GE were reviewed carefully by the Kuosheng senior reactor operator (SRO) and EOP expert, the online data was verified by running the simulator with ERFTDS and using a Kuosheng developed software to monitor and dump the data from the EOPTS database. The problems experienced during this phase could be classified as follows:

a. The simulator data point selection was incorrect.
b. The engineering unit conversion error.
c. The compose point algorithm of the GE database was incorrect.
d. The data point definition in the EOPTS database dictionary was incorrect.
e. The GE database was insufficient for the EOPTS.

The incorrect compose points algorithm and data point definition were modified and the insufficient database were added then feed back to GE.

2.6 EOPTS RULES VERIFICATION

The EOPTS rules include the following:

a. General Control (GENCTL.RUL)
b. Reactor Pressure Vessel Control (RPVCTL.RUL)
c. Primary Containment Control (PCCTL.RUL)
d. Secondary Containment Control (SCCTL.RUL)
e. Radioactivity Release Control (RRCTL.RUL)
f. Contingencies Control (CONTCTL.RUL)

The EOPTS rules verification were performed by insert malfunction(s) to the simu -lator to create the EOP entry condition(s), then froze the simulator to verify that the appropriate emergency operating procedure(s) were entered, the EOP step and messages were correctly displayed on the VT220 screen and none conflict messages were displayed on the screen at same time, then run the simulator again. If any error was found, the associated EOPTS rule logic and/or database should be rechecked, corrected and retested until it was satisfactory.

There were numerous questions of the EPGs had discovered during the EOPTS rules verification (see ATTACHMENT), it should be clarified and/or specified by the

BWR Owner Group or somebody else, then the EOPTS could be exactly prepared to be used in the BWR nuclear power plant.

2.7 MAN-MECHINE INTERFACE TURN UP

The operator comments for the EOPTS were as follows:

a. The response time of the ASK-USER was too long.
b. The EOP messages were erased and then refreshed too fast.
c. The screen manager was died sometimes, when the SEE_MORE function being in use.

The screen manager was modified to response the ASK_USER immediately after the operator key in. The screen manager code was changed to erase the out of date message and insert the new message only, for operator easy to read, To send the SEE_MORE messages line by line, instead of directly %S format, to prevent it to die.

2.8 SIMULATOR MODEL MODIFICATION

The simulator model was limited so that it was not feasible to run all the EOP's scenarios, the database may not enough for used in the EOPTS and the simulator was gone crazy (computer hung up) sometimes during a severe transient.

The simulator database were added when necessary, the simulation model were modi -fied or added to provide the feasibility to run the most EOP's scenarios and some limits such as rate of change of the reactor water mass inventory, reactor core moderator quality which shall not be negative, any equation shall not be zero divided by a parameter, etc. were added to prevent the computer from hanging up.

SECTION 3

SUMMARY OF THE PROJECT

The project had complected at Feb. 25, 1989, after the EOPTS evaluated by all of the Kuosheng main control room operator shift crews (6 shift groups split into 12 crews). The implementation of the Emergency Operating Procedure Tracking System in the Kuosheng Simulator, Taiwan Power Company have gained the following benefits:

a. Gained the high technology of Artificial Intelligent System.
b. Improved the Kuosheng simulating functions.
c. Gained a very effectual tool to verify and validate the REFTDS as well as the SPDS via the ERFTDS simulation.
d. Gained the technology of development and modification of the EOPTS logic and rules.
e. Verified and validated the Kuosheng Emergency Operating Procedures.
f. Gained a very effectual simulator for operator training of the ERFTDS, SPDS and EOPs.
g. Provided a good facility for plant emergency drill.

ATTACHMENT

SUBJECT: QUESTIONS OF THE EMERGENCY PROCEDURE GUIDELINES FOR PREPARE THE KUO-SHENG EOP TRACKING SYSTEM

The following Quations of emergency procedure guidelines were discovered during the implementing Kuosheng EOP tracking system, it shall be clarified and/or specified by the BWR Owner Group or somebody else, then the EOP tracking system could be exactly prepared to be used in the BWR nuclear power plant.

REFERENCES:
1. BWR OWNERS' GROUP EMERGENCY PROCEDURE GUIDELINES OEI Document 8390-4, Draft Revision 4AF, August 14, 1986
2. MARK III CONTAINMENT HYDROGEN CONTROL SUPPLEMENT Draft Revision 4AB, October 31, 1985

1. How to decleare that it is "Cannot be Determined" ? It should be to listed all plant available indications related to it in the EOPs.

   Example: RPV water level "cannot be determined", enter [procedure developed from .....].

2. How to determined that "The Reactor Will Remain Shutdown Under All Condition Without Boron" ? Is it determined by the nuclear engineer or by the reactor operator ? What is the time limit for them to determined it ?

   Example: Any control rod cannot be inserted to .... and it has not been determined that "the reactor will remain shutdown under all conditions without boron', enter [.....].

3. How far "Before" the identified parameter to reaches a limit or action level then the operator shall take the specified action ?

   Example: "Before" suppression pool temperature reaches [the Boron Injection Initiation Temperature] ..... then ....

4. When should the operator be initiated the SBLC ? Since reactor power may be oscillating up and down due to RPV water level increase or decrease, "BORON INJECTION IS REQUIRED" may comes to TRUE then FALSE.

   Example: Before suppression pool temperature reaches [the Boron Injection Initiation Temperature] but only if the reactor cannot be shutdown "BORON INJECTION IS REQUIRED", inject boron into the RPV ......

5. What is the margin and time limit (from reaching the margin to the limit or action level, i.e., decreasing or increasing rate) of the identified parameter for operator to determined that it "Cannot be Maintained Above (or Below)" the specified limit or action level ?

   Example: If primary containment water level "cannot be maintained below" the Maximum Primary Containment Water Level Limit, terminate injec-tion into the RPV .....

6. What is the time limit of the identified parameter not return to and remain above (or below) the specified limit or action level, then said that it "Can not be Restored and Maintained Above (or Below)" the specified limit or action level ?

   Example If drywell or suppression chamber (containment) hydrogen concentration "cannot be restored and maintained below" 6%, then ....

7. What is the definition of "SRV is Cycling" (i.e., the time limit of a SRV from closing to reopen) ? If any SRV is cycling on Low Low Setpoint logic (BWR-6 design), should the operator need to manually open the SRVs until RPV pressure drops to [.....] ?

   Example: If any "SRV is cycling", initiate IC and manually open SRVs until RPV pressure drops to [935 psig (RPV pressure at which all.....)].

8. How long (time limit) from the specified condition(s) are meet to the time the action cannot be accomplished then said it "Cannot be ..." ?

   Example: When the shutdown cooling RPV pressure interlock clears, initiate shutdown cooling ..... If shutdown cooling "cannot be established" and .....

9. What is the definition of "Further Cooldown is Required" (i.e., under what condition(s) further cooldown is required) ?

   Example: If shutdown cooling cannot be established and "further cooldown is required", continue to cool down using .....

10. Should the operator need to check the RPV water level is above the TAF or not, before they take the action of "Prevent Automatic Initiation of ADS" ?

    Example: Before suppression temperature reaches ... ; inject boron into the RPV with SBLC and "prevent automatic initiation of ADS".

11. When suppression pool temperature cannot be maintained below the Heat Capacity Temperature Limit. Why not lower the RPV pressure to below the HCTL first ? (refer to page RC-9). Suggest change SP/T-3 and add SP/T-4 to read as follows:

    SP/T-3 If suppression pool temperature cannot be maintained below the Heat Capacity Temperature Limit,maintain the RPV pressure the below the limit, enter [procedure developed from the RPV Control Guidelines] at [Step RC-1] and execute it concurrently with this procedure.

    SP/T-4 When suppression pool temperature and RPV pressure cannot be maintained below the Heat Capacity Temperature Limit, EMERGENCY RPV DEPRESSURIZATION IS REQUIRED.

12. When suppression pool water level cannot be maintained above the Heat Capacity Level Limit, why not lower the RPV pressure to above the Limit first ? Since lower the RPV pressure will increase the Heat Capacity Temperature Limit, results Heat Capacity Temperature Difference increase and Heat Capacity Level Limit decrease, Suggest change SP/L-2,1 to read as follow:

SP/L-2,1 Maintain suppression pool water level above the Heat Capacity Level Limit.

....

If suppression pool water level cannot be maintained above the Heat Capacity Level Limit, lower the RPV pressure to above the Limit, enter [procedure developed from the RPV Control Guidelines] at [Step RC-1] and execute it concurrently with this procedure.
If suppression pool water level and RPV pressure cannot be maintained above the Heat Capacity Level Limit, EMERGENCY RPV DEPRESSURIZATION IS REQUIRED.

13. When primary containment water level cannot be maintained below the maximum Primary Containment Water Level Limit, should lower the suppression chamber (containment) pressure to below the Limit first (refer to page RC-3), Suggest change SP/L-3,3 to read as follow:

SP/L-3,3 Maintain primary containment water level below the Maximum Primary Containment Water Level Limit.

....

If primary containment water level cannot be maintained below the Maximum Primary Containment Water Level Limit, then irrespective of the offsite radioactivity release rate , vent the primary containment, defeating isolation interlocks if necessary, to reduce and maintain the suppression chamber (containment) pressure to below the Limit.
If primary containment water level and suppression chamber (containment) pressure cannot be maintained below the Maximum Primary Containment Water Level Limit, then irrespective of whether adequate core cooling is assured terminate injection into the RPV from source external to the primary containment until primary containment water level and suppressure chamber (containment) pressure can be maintained below the Limit.

14. Should the operator need to check that there is any system, injection subsystem or alternate injection subsystem is line up with at lwast one pump running or not, before they take the action of "EMERGENCY RPV DEPRESSURIZATION IS REQUIRED" ?

What should the operator do, if no system, injection subsystem or alternate injection subsystem is available and EMERGENCY RPV DEPRESSURIZATION IS REQUIRED ?

When is the emergency RPV depressurization complected ? When the condition of EMERGENCY RPV DEPRESSURIZATION IS REQUIRED clears or RPV has depressurized to less than 050 psig (Minimum SRV Reopening Pressure) above suppression chamber (containment) pressure] ?

Example: When drywell temperature cannot be maintained below [340 F (maximum temperature at which ADS .... (], "EMERGENCY RPV DEPRESSURIZATION IS REQUIRED", enter [procedure ....

15. Why not continue operate the drywell hydrogen mixing system, if drywell hydrogen concentration is reaches 6% but containment hydrogen is below 6% ? Since drywell hydrogen mixing system is take suction from the containment

(low H2 concentration) discharge to the drywell (high H2 concentration) then push the vapor and gas in the drywell (high H2 concentration) thru suppression pool horizontal vents to the containment (low H2 concentration) to reduce the drywell H2 concentration.

Example: [When drywell or suppression chamber hydrogen concentration reaches 6%], EMERGENCY RPV DEPRESSURIZATION IS REQUIRED; .... "secure hydrogen mixing system" and .....

16. Why [RPV pressure is below the Primary Containment Pressure Limit] is one of the conditions for drywell hydrogen mixing system operation ? Since RPV pressure is nothing to do with the Primary Containment Pressure Limit, Even the primary containment pressure will not affect by the operating of drywell hydrogen mixing system, the drywell hydrogen mixing system is take suction from the containment and discharge to the drywell, then push the vapor and gas in the drywell through the suppression pool horizontal vents back to the containment.

Example: Before drywell hydrogen concentration reaches [4% (lowest hydrogen concentration ....)] but only if "[RPV pressure is below the Primary Containment Pressure Limit and]" drywell and suppression chamber hydrogen concentration are below 6 %, operate the drywell hydrogen mixing system.

17. Does the following emergency procedure guidelines override the radioactivity release control guideline RR-1 or not ?

PC/P-4 Before suppression chamber (containment) pressure reaches [the Primary Containment Pressure Limit], then irrespective of the offsite radioactivity release rate, vent the primary containment, ...

PC/H If while executing the following steps:

Drywell or suppression chamber (containment) hydrogen concentration cannot be determined to be below 6%, EMERGENCY RPV DEPRESSURIZATION IS REQUIRED; enter ....; " irrespective of the offsite radioactivity release rate " vent and purge primary containment .....

PC/H-4 [When drywell or suppression chamber (containment) hydrogen concentration reaches 6%], EMERGENCY RPV DEPRESSURIZATION IS REQUIRED; enter ....; secure hydrogen mixing system and, " irrespective of the offsite radioactivity release rate " vent and purge primary containment .....

C6-3 When primary containment water level reaches [26 ft 3 in. (elevation of ....)], then "irrespective of the offsite radioactivity release rate" vent the RPV, defeating .....

18. What is the time limit for operator to line up injection subsystems and alternate injection subsystems, before they take the next action ?

Example: When RPV water level drops to [......(top of active fuel)], ... If any system, injection subsystem or alternate injection subsystem is line up with ..... then ......

If no system, injection subsystem or alternate injection subsystem is line up with .... then ......

19. Should the operator take the action of EMERGENCY RPV DEPRESSURIZATION IS REQUIRED, if only one ECCS keep-full systems, SLC (test tank), or SBLC (boron tank) alternate injection subsystem is line up with at least one pump running ? (i.e., dose the RPV will be able to get Adequate Core Cooling after emergency RPV depressurization, by one of such a small capacity alternate injection subsystem ?)

Example: When RPV water level drops to [......(top of active fuel)], ... If any system, injection subsystem or alternate injection subsystem is line up with at least one pump running, EMERGENCY RPV DEPRESSURIZATION IS REQUIRED.

20. How to performing the Emergency RPV Depressurization, if suppression pool water level is below [4 ft 9 in (elevation of top of SRV discharge device)]?

CS-1,3 If suppression pool water level is above [4 ft 9 in. (elevation of top of SRV discharge device)]:

* Open all ADS valves.
* If any ADS valves cannot be opened, open .....

Suggest change C2-1.4 to read as following:

C2-1.4 If suppression pool water level is below [4 ft 9 in. (elevation of top of SRV discharge device)] or less than [3 (Minimum Number of SRVs Required for Emergency Depressurization)] SRVs are open [and .....], rapidly depressurize the RPV ....

21. How to performing the "Steam Cooling" for a plant did not has the IC ?

C3-1 Confirm initiation of IC.
....

22. What should the operator do, after RPV flooding to EPG's step C4-1.4 but not all control rods can be inserted to or beyond position [02 (Maximum Subcritical Banked Withdrawal Position)] and it has not been determined that the reactor will remain shutdown under all conditions without boron ? Since if the operator continue injecting boron with SBLC or alternate boron injection system, the reactor power and pressure will decrease, operator will increase injection to maintain at least [1 (minimum number of SRVs ....)] SRV[s] open and RPV pressure above the Minimum Alternate RPV Flooding Pressure, eventual -ly will flooding the RPV to above MSL and discharge the reactor water with boron thru SRVs to the suppression pool.

23. At what step should the operator be " Continued in this procedure " of the following EPGs ?

Example: Terminate and prevent all inject ..... until RPV pressure is below .....

If less then [1 (minimum number of SRVs for ......)] SRV[s] can be opened, "continue in this procedure". (C4-1.1, C5-3.1)

C4-1.5 When all control rods are inserted to or beyond position [02 (maximum Subcritical Banked Withdrawal Position)] or it has been determined that ....., "continue in this procedure".

Suggest change C4-1.5 to read:

C4-1.5 When all control rods are inserted to or beyond position [02 (Maximum Subcritical Banked Withdrawal Position)] or it has been ....., "continue in this procedure at [Step C4-3]".

24. Why the condition(s) for isolate steam lines are differente for case of All Rods In and Not All Rods In ?

C4-1.2 If at least [3 (Minimum Number of SRVs Required for Emergency Depressurization)] can be opened, close the MSIVs, main steam line drain valves, and IC, RCIC, and RHR steam condensing isolation valves.

C4-2 If at least [3 (Minimum Number of SRVs Required for Emergency Depressurization)] can be opened, close the MSIVs, main steam line drain valves, and IC, RCIC, and RHR steam condensing isolation valves.

25. How to get the RPV pressure to below the Minimum Alternate RPV Flooding Pres -sure after terminate and prevent all injection into the RPV except from boron injection systems and CRD ?

Example: Terminate and prevent all injection into the RPV except from boron injection system and CRD "until" RPV pressure is below the Minimum Alternate RPV Flooding Pressure.
.... (C4-1.1, C5-3.1)

26. Is it feasible to change C4-1.1, C5-3.1, C4-1.2 to read as following ? Since terminate and prevent all injection into the RPV and RPV emergency depressurization should be performed in Contingency #2 (refer to pages C2-2, RC-8)

C4-1.1 Continue in [procedure developed from the Contingency #2] at [Step C201.3] or [Step C2-1.4] until RPV pressure is below the minimum Alternate RPV Flooding Pressure.
....

If less than [1 (minimum number of SRVs for with the ....)] SRV[s] can be opened, continue in this procedure at [Step C4-1.3].

C4-1.2 When RPV pressure is emergency depressurized to below the Minimum Alternate RPV Flooding Pressure, close the MSIVs, main steam line drain valves, and IC, RCIC, and RHR steam condensing isolation valves. (i.e.; isolate the steam lines for easy to flooding the RPV to above the Minimum Alternate RPV Flooding Pressure, after emergency RPV depressurization is done.)

27. When should the operator commence and increase injection into the RPV for RPV flooding ?

Example: Commence and, ... increase injection into the RPV with the following systems until ......

28. What is the time limit for operator to try every efforts and then judgment that the first action "Cannot be Accomplished" then take the next action ?

C4-1.3 Commence and .... until at least [1 (minimum number of... ....(] SRV[s] [is] open and RPV pressure is above the Minimum Alternate RPV Flooding Pressure:

If less than [1 (minimum number of ......)] SRV[s] [is] open or RPV pressure "cannot be increased to above the Minimum Alternate RPV Flooding Pressure", commence and ....

If less than [1 (minimum number of ......)] SRV[s] [is] open or RPV pressure "cannot be increased to above the Minimum Alternate RPV Flooding Pressure", enter [procedure developed from Contingency #6] and ....

29. How to get adequate core cooling during RPV Flooding, when commence injection at the time the RPV pressure is below the Minimum Alternate RPV Flooding Pressure but above the shut off head (i.e., no injection flow) of all available injection system(s) ? (especially in case of 1 or 2 or no SRV(s) can be opened)

30. Are the operator allowed to close the SRV(s) to increase the RPV pressure to above the Minimum Alternate RPV Flooding Pressure (but below the shut off head of the available injection system(s)) and keep at least [1 (minimum number of SRVs for which the Minimum Alternate RPV Flooding Pressure is below the lowest SRV lifting pressure)] SRV[s] [is] open to prevent enter [procedure developed from Contingency #6] ?

i.e., Change "...... SRV[s] [is] open or RPV pressure cannot be ...."
to "...... SRV[s] [is] open and RPV pressure cannot be ...."
in the EPG C4-1.3

31. Are the operator allowed to close the SRV(s) to maintain the RPV pressure to at least [75 psig (Minimum RPV Flooding Pressure)] above suppression chamber pressure and keep at least [3) Minimum Number of SRVs Required for Emergency Depressurization)] SRV[s] are open to prevent enter [procedure developed from Contingency #6] ?

i.e., Chang "...... SRV[s] are open or RPV pressure cannot be ...."
to "...... SRV[s] are open and RPV pressure cannot be ...."
in the EPG C4-3.1

32. Does enter [procedure developed from Contingency #6] is the only way for RPV Flooding, if less than [1 (minimum number of SRVs for ....)] SRVs can be opened or less than [3 (Minimum Number of SRVs Required for ....)] SRVs can be opened in Contingency #2 ?
(i.e., How to accomplished the RPV Flooding, if either of the above case is existed ?)

Reference to the EPGs C4-1.3, C4-3.1 and C4-4

A Knowledge-based System for PWR Loading Pattern Determination

PASCAL DAUBOIN
Electricite de France Research Center
Information Processing and Mathematical Studies Department
1, Avenue du General de Gaulle
92141 Clamart, Cedex, France

ABSTRACT

This paper deals with the design of a knowledge based system for solving of an industrial problem which occurs in nuclear fuel management. The problem lies in determining satisfactory loading patterns for nuclear plants. Its primary feature consists in the huge search space involved. Conventional resolution processes are formally defined and analyzed: there is no general algorithm which guarantees to always provide a reasonable solution in each situation. We propose a new approach to solve this constrained search problem using domain-specific knowledge and general constraint-based heuristics. During a preprocessing step, a problem dependent search algorithm is designed. This procedure is then automatically implemented in FORTRAN. The generated routines have proved to be very efficient in finding solutions which could not have been provided using logic programming. A prototype expert system has already been applied to actual reload pattern searches. While combining efficiency and flexibility, this knowledge based system enables human experts to rapidly match new constraints and requirements.

INTRODUCTION

The problem we address here is to determine the correct reload pattern for fuel assemblies in a nuclear plant. All nuclear reactors must usually be reloaded once a year. Satisfactory locations for assemblies have to be chosen within the core. The power distribution of a successful configuration is required to meet safety specifications.

Nuclear plant loading pattern design is an extremely significant real case of combinatorial problem. Assuming that the n assemblies to be reloaded in a n-element nuclear core have previously been selected, the number of repositioning matrixes liable to be produced (**M(n)**) is obtained using the following formula:

$$M(n) = (n!) * r^n$$

where **r** is the number of possible rotations applicable to the assemblies.

A 900-M.W. P.W.R. reactor core which includes 157 assemblies is shown on Figure 1.

The standard strategy currently adopted by Electricité de France prevents assembly rotations on site. Moreover, new fuel assemblies have a preset position (they are placed at the core periphery). Ultimately, the number of possible rearrangements is about (100!). Obviously, a blind search of this state space cannot be performed.

CONVENTIONAL METHODS OF SOLUTION

On the one hand, the conventional solution relies on the "trial and error" paradigm: human experts shuffle the assemblies, evaluate the candidate configuration with a mainframe-based program, analyze the output , generate a new configuration and repeat the process until a good solution is reached. The evaluation routine included in this iteration loop is extremely time consuming. Ordinarily, experts try to recognize a familiar core situation which leads to plausible arrangements. However using previous results of analogous situations becomes less and less tractable because of a plant's singular history (more and more irregularities exist among assemblies).

On the other hand, several optimization methods have been proposed either to minimize the unit fuel cost, or to maximize safety margins (8,10,13). Based on small perturbation theory, this approach seems to be less empirical than the former one. But these procedures usually need a reference loading pattern as a starting point. As this initial step still has to be performed manually, it encounters the same problem as the forementioned strategy. Furthemore, in numerous instances the changes due to assembly shuffling can have far reaching effects and they are not small perturbations.

Although it is possible to make use of a "brute force" technique for partial exploration of the problem raised, this sole development line does not meet the time requirement. The computation time varies exponentially with the problem size and quickly becomes prohibitive. There is no general algorithm which guarantees to always provide a reasonable solution to each core situation. Thus, great attention has been paid to the potential use of A.I. tools.

A SECOND GENERATION EXPERT-SYSTEM

Combinatorial analysis thus compels the use of domain knowledge. Some systems try to do so using repositioning matrixes set by experts (7,10). However, knowledge is, in this case, expressed under compiled form. Indeed, a whole range of prior exploration work on the possible arbitrations among various alternatives, and of compromise among various constraints is thus bypassed and only the end result of this decision making process is retained. Shallow reasoning (in that a large part of the expert work does not appear) does not allow the systems resorting to such knowledge to modify their strategy in the event of a deadlock. Such processes can therefore only handle a limited number of problem instances.

In the proposed approach, we intend to model the underlying cognitive processes in order to recognize and rebuild the principles which have enabled human experts to become actual skilled experts. Besides, the in-depth explanation of the human strategy makes it possible to consider domain knowledge as explicit objects on which we can apply new knowledge (meta-knowledge). Moreover, it must be pointed out that the nuclear fuel management is an ever-changing technique, both at the technological level (assembly modification) and at the economic level (management matching the network demand for instance).

We have therefore adopted a declarative approach, separating inasmuch as possible, the solution requirements from how the work is to be carried out. In this way, constraint specification represents a convenient form for stating what kind of configurations must be achieved, turning more of our attention towards the description of the target.

Much of the design process of a loading pattern depends on recognizing, formulating and satisfying these constraints. Dealing with the latter constraints in which form, function and physics strongly interact is a difficult task. These conditions are well suited to the use of Knowledge Based Systems.

As an initial step towards the acquisition of deep knowledge, a model has been developed to determine loading patterns in P.W.R. focusing on the reactivity distribution. The problem consists in assigning values (assemblies to be loaded into the core) to variables (locations within the core) which are subject to a set of constraints (technical limitations and specifications for assembly shuffling).

Methodology

Our purpose is to determine whether the prototype knowledge based system design meets certain specification contraints (e.g., power of expression, flexibility, response time).

As shown on figure 2, the method of solution is subdivided into two parts. First, given the problem statement, a strategy for efficiently searching the branching tree of the possible loading patterns is determined.

This preprocessing step defines a problem dependent algorithm scheme which is oriented to find a single solution (the first one). Secondly the search procedure is automatically implemented in an efficient language programming (namely FORTRAN) so that a practical solution may be obtained within a reasonable response time.

When the generated routine is run, it outputs a satisfactory loading pattern, otherwise the problem data have proved to be not suitable to fulfil the requirements (see fig. 2).

Defining an "ad hoc" search algorithm

As it can be noticed from figure 2, a Knowledge Base is used to design the search algorithm prior to running the exploration of possibilities. It is made up of two parts: a general purpose subsystem gathering constraint-based heuristics and a production rule subsystem which includes domain specific knowledge. The latter is activated at the begining of the resolution performing three main functions:

i/ Problem specification in terms of domain variables, values that should be assigned and constraints between variables. Note that predicate calculus features allow adequate statement of generic principles such as symmetry constraint in this rule base. These principles in turn lead to instantiated constraints which apply on the particular problem instance. A constraint is said to be instantiated when the variables which are involved in its definition are bounded to objects in the domain. Here is a production rule according to which every pair of symmetrical locations must receive assemblies with similar physical characteristics :

IF

(L1)	is_a	location
(L2)	is_a	location
(L1)	symmetrical	(L2)
(L1)	possible_instance	(A1)
(L2)	possible_instance	(A2)
(F)	is_a	physics_function

THEN

ABS((F)(A2) - (F)(A1)) less_than ε

where (L1), (L2), (A1), (A2), (F) are production system variables.

This generic constraint implicitly represents more than 1000 numerical constraints for a complete core. As can easily be noticed, the problem statement is greatly simplified by logical variables and relational forms which allow easy handling of a variety of formulations.

ii/ Early pruning to limit the combinatorial explosion. A set of shuffling rules and basic heuristics greatly reduces the number of a priori possible configurations. They focus on specified limitations (which deal with fresh assemblies, control rods, locations on axis among others) in order to prevent useless exploration of alternatives. Let us take a straightforward example. The following restraint must apply : locations placed beneath a control rod should house assemblies with low reactivity. The corresponding rule is written as follows :

IF

(L)	is_a	location
(L)	is_under	(CR)
(CR)	is_a	control_rod
(A)	is_a	assembly
(RA)	reactivity_of	(A)
(RA)	greater_than	low_reactivity_level

THEN

REMOVE ((L) possible_instance (A) *)* iii/ Correct value ordering. When instances compatible with a variable cannot be positively discarded (previous task), it is sometimes possible to generate a priority order for assignment of fuel elements to preset core locations. For instance, it is advised to relocate on symmetry lines assemblies which were placed on these lines over a previous cycle. Such rules provide a static order of values to be assigned to variables. However, when an evaluation function that can discriminate the candidate values for a variable is available (this function usually depends on previously assigned variables), it can be safely incorporated into the search algorithm. During the exploration of possibilities, for example, a checkerboard pattern of high and low reactivity assemblies is sought. This is performed with a view to achieving a flat power distribution. Hence, every element selected for a given location influences the future assignments of its neighbouring locations.

In both cases (static or dynamic selection), the value order may be obtained by symbolic or numerical means resulting in a partial or exhaustive classification. When such guidelines are taken into account, it is possible, at decision tree path level, to start by selecting one element rather than another for a given variable.

These inferences are driven by the problem instance data and end up with a complete definition of the underlying constraint network. Regardless of the application dependent strategies, a second rule based subsystem uses the variable dependencies from the problem constraint network to select an efficient order by which variables get instantiated. Studies on constrained search problems (4,5,11) have shown how the variable order has a tremendous effect on the exploration procedure's performances since each ordering defines a different search space with a different size. Hence, an evaluation function is computed to find out how each variable constrains the rest of the search space. Each variable is given a rank which depends on the number of corresponding possible values and on the number (and nature) of constraints where it participates.

The suggested method considers a predetermined ordering which cannot vary dynamically during the search (3,12). According to this variable order, constraints are posted in the algorithm so as to be checked as soon as possible during execution. This is intended to prune the search space in the most effective way.

Automatic programming.

The solution space can be expressed as a tree structure in which each node corresponds to the assignment of a variable by a certain value. Once the Knowledge Base has proceeded through all deductions, an efficient "top-down" procedure for the exploration of the branching tree is determined (i.e. a variable ordering, the subsequent constraint posting, and a partial value order).

This forward search needs a backtracking procedure to go backwards when a dead-end occurs (i.e. when all possible values for a given variable have been tried without success). Although selective backtracking substantially reduces the backtracking effort since it consists in returning to the failure source, only a chronological backtracking has been applied at the current stage of development.

These forward and backward procedures must be recursively applied until a solution is reached. The search algorithm is now thoroughly defined. Hence, it is possible to automatically generate an implemented code that matches this predetermined scheme.

The underlying ground for automatic programming is the use of an efficient conventional language (such as FORTRAN,C,PASCAL ...) to find solutions which could not have been provided using logic programming. Furthermore, this program synthesis step relieves the user from tree search programming.

For testing purposes, the generated codes are written in FORTRAN. It should be noted that the generated program greatly depends on the problem structure but also on the numerical data. Each problem instance leads to a particular routine adapted to the treatment of its own search space.

Nevertheless, generated FORTRAN routines can include parameters matching the special demands of domain experts. Given a constraint, the corresponding threshold can be treated as a variable during search algorithm determination. Chosen values are assigned to parameters before running the exploration code.

Owing to this feature, the same generated routine can be reused for new requirements provided that the constraint network structure remains the same. For example, when the requirements are so tight that no solution is obtained, constraint limits may be adjusted. More generally, tradeoffs between specifications are often necessary so as to provide judicious fuel element arrangements.

IMPLEMENTATION AND RESULTS

The global system has already been applied to actual reload pattern searches with real plant data (under equilibrium conditions). Nuclear core configurations have been generated on a quater core basis (1,3).

The results are related to a standard fuel management program: "out-in" three region cycling. For this application, a forward chaining inference engine based on first order logic : Genesia II is used (6,9). The characteristics of the problem are set into a factual base (about 1000 facts are necessary to describe the fuel management scheme and the selected assembly characteristics). Domain specific knowledge is given in an explicit declarative form amouting to about 50 rules which are based on predicate calculus. More than 300 specific constraints are derived from these basic principles. The constraint reasoning component is made up of 200 first order rules and the FORTRAN implementation task is achieved by means of about 40 rules.

The average time for search procedure generation is around 2 minutes (on an IBM 3090 MVS/XA), including automatic FORTRAN implementation. The response time slightly varies with the size of the constraint network.

Alternative feasible solutions have been examined providing loading pattern with different features (dealing with core symmetry or assembly corner adjustment for instance).

Despite the fact that the Knowledge Based system does not make any attempt to optimize the solution, parameters have easily been modified in order to refine the current solution. Successful core configurations have been generated within satisfactory response time (ranging from 0.005 to 0.8 s).

CONCLUSIONS

This paper discusses a new approach to nuclear plant loading pattern determination. The method of solution makes use of domain-independent techniques (constraint reasoning and program synthesis) as well as domain specific knowledge. It stems from the first results that the approach presented here can be extended to new kinds of in-core fuel management. Although the problem faced is highly combinatorial, the average behavior of the predetermined search procedures has proved to be very satisfactory. The method of solution is significantly improved by matching the structure and data of the particular problem to be solved. While combining efficiency (due to the problem oriented resolution) and modularity (due to the declarative nature of the knowledge involved), this Knowledge Based system enables human experts to rapidly check new constraints and strategies.

REFERENCES.

1. DAUBOIN P.and LUCAS J.Y."An automatic nuclear core configuration generator" . Proccedings of the international workshop on artificial intelligence for industrial applications . IEEE AI'88. Hitachi, Japan (1988).

2. DAUBOIN P. , "A Knowledge Based System for Nuclear Plant Loading Pattern Determination, " Proceedings of the AAAI Spring Symposium, Stanford, U.S.A. (1989).

3. DAUBOIN P. , "Un Système-Expert pour la recherche de plan de rechargements" Note EDF/DER HI/ 21-6348 (1989).

4. DECHTER R. and PEARL J. "Network-Based Heuristics for Constraint Satisfaction Problems," Artificial Intelligence Vol. 34 (1988).

5. De KLEER J. "An assumption-based TMS". Artificial Intelligence Vol. 28 n° 2 (1986).

6. DORMOY J.L. "Résolution qualitative: complétude, interprétation physique et contrôle. Mise en oeuvre dans un langage à base de règles: BOOJUM" . Thèse de l'Université Paris 6, (1987).

7. FAUGHT W.S. and ROTHLEDER B.M., "A Prototype Fuel Shuffling System Using a Knowledge Based Toolkit," Computer applications for nuclear plant operation and control, American Nuclear Society Meeting, Pasco,WA (1985).

8. FEDEROWICZ A.J. and STOVER R.L., "Optimization of PWR Loading Patterns by Reference Design Perturbations ," Trans. Am. Nucl. Soc., 17, 308 (1973).

9. GENESIA II : Générateur de Systèmes-Experts. Manuel de référence version 1.1 Steria Group (1988).

10. HOBSON G. H. and TURINSKY P. J. , "Automatic Determination of Pressurized Water Reactor Core Loading Patterns that Maximize Beginning-of-Cycle Reactivity whithin Power-Peaking and Burnup Constraints," Nucl. Technol., 74, 5 (1986).

11. LAURIERE J.L. "A language and a program for stating and solving combinatorial problems". Artificial Intelligence Vol. 10 (1978).

12. LUCAS J.Y " SIREN, un résolveur général de problèmes combinatoires," Note EDF-DER HI/21-5761-02 (1987).

13. MOTODA H. , HERCZEG J. , and SESONSKE A. , "Optimization of Refueling Schedule for Light-Water Reactors," Nucl.Technol., 25, 477 (1975).

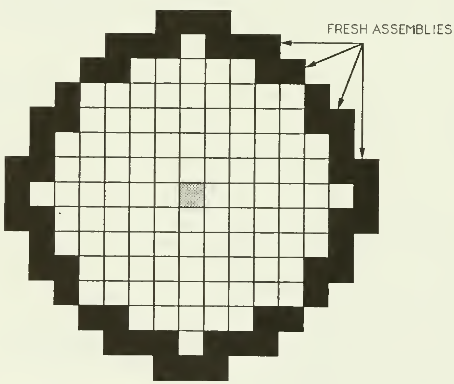


Figure 1. Topography of a 900-MW P.W.R Core

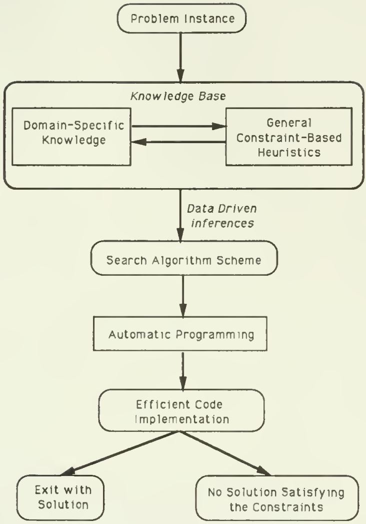


Figure 2. System's Framework.

An AI-based Planning System for Core Shuffles

CARL H. NEUSCHAEFER
Combustion Engineering, Inc.
NPBU, Nuclear Services
Windsor, Connecticut, USA

STEFAN GONICK
Combustion Engineering, Inc.
NPBU, Nuclear Services
Windsor, Connecticut, USA

JOSEPH NASER
Electric Power Research Institute
3412 Hillview Avenue
Palo Alto, California 94304, USA

ABSTRACT

In preparation for a refueling outage and during the outage itself, utility personnel become concerned with the generation and monitoring of a crane/fuel movement sequence (core shuffle plan). The core shuffle plan is the sequence of steps involving the movement and placement of core components for refueling purposes. Given an initial (existing) core configuration, a final (core reload) core configuration and plant conditions and equipment, the planner determines the core shuffle plan. The planning process becomes more involved and important when one considers: minimizing crew and/or outage time; minimizing tool changes; constraints on fuel, control rod support, or refueling mast orientations, etc.; and the particular plant equipment available at the start (let alone should it change during the outage). Further, the ability to monitor the execution of the plan i.e. to track and accurately maintain a status and record during the course of the outage and to support replanning when problems are encountered are significant. Several efforts have been made to explore automating the process of plan generation. None to date have completely addressed the generic needs.

This paper describes the results of an EPRI project performed by Combustion Engineering, Inc., Nuclear Services to develop a more encompassing and

flexible computer based core shuffle planning system. A system which provides the extensive planning and monitoring capabilities needed. The software developed is based on a combination of traditional software procedural methods with enhancements incorporated readily with certain Artificial Intelligence (AI) software techniques. These enhancements along with the core shuffle planning system functionality are described.

I. INTRODUCTION

Some effort has been spent on the part of various organizations to develop planning systems for core shuffles (References 1-4). A full-scale insert shuffle planning system prototype has been developed by EPRI for the case of a PWR where the core is totally off-loaded into the spent fuel pool and the inserts are shuffled there. Combustion Engineering, Inc., (C-E) had a nuclear fuel shuffling sequencer, which generates a shuffle sequence based upon minimizing the time/distance of refueling machine travel. The refueling sequence can be generated for a normal, over-the-core shuffle in PWR's. Neither the prototype system developed by EPRI nor the original sequencer developed by C-E is general enough to handle the full scale problem of shuffling fuel assemblies and inserts, either inside the core or in the spent fuel pool or monitoring shuffle plan execution. Also, the two systems had only addressed the problem from the PWR utilities' point of view. This paper describes the results of an effort to develop a more general and comprehensive system for both PWR's and BWR's. The system incorporates traditional software techniques with some Artificial Intelligence (AI) techniques to enhance the functionality.

The manual development of a crane movement sequence for fuel and insert shuffling requires extensive engineering time (two to four man-weeks). Further, the ability to review and validate and/or to make changes to a plan during an outage evolution are time critical. Due to the length of time to manually develop and/or modify and verify a shuffle plan, it is frequently not possible to look at alternative strategies which could lead to a more effective or efficient (less time required) shuffle sequence. EPRI, as a result of previous work (Reference 1), has established that an expert system approach could develop efficient shuffle plans and allow modifications to the plans quickly, to reduce the considerable man-power and time (planning and outage) currently expended. EPRI has sponsored an expert system software implementation project to develop a generic fuel shuffle planning system.

The result of this project is a system intended to be used by the PWR and BWR utility engineers currently involved in generating shuffle plans, and by the engineers and crane operators who execute those plans. The purpose of this system is to produce complete plans for the shuffling of fuel from an initial core configuration to a desired reload core configuration for three cases: 1) PWR in-core shuffles, 2) PWR off-load/reload core shuffles, and 3) BWR in-core shuffles. An automated system would: reduce outage time thru efficient plans; reduce manhour costs to prepare plans and reduce time and effort to modify plans (particularly during critical outage situations); perform extensive error checking and validation; and allow for on-line monitoring and tracking of the execution of the plan during the outage for rapid and accurate status and record generation.

The shuffle planning system has been designed on a P.C. class workstation utilizing an expert system software architecture. The system provides a modularized software design to provide the shuffle planning and user interface functionality. The system automates the process of creating fuel shuffle plans with the attending information and decision computer support aides, providing a sophisticated yet simple to use interactive planning workstation. A window and menu oriented user interface guides the user thru initial setup, planning, verification and report generation. A software interface exists to allow access to external database information (such as a Nuclear Fuel Accountability System). The software is written in LISP and utilizes an object-like data structure. The following sections will provide more detail and insight into the design approach and its implementation features.

II. CORE DESIGN AND SHUFFLE BACKGROUND

Light Water Reactor(s) (LWR) are required to be shut down periodically for replacement of expended fuel assemblies. The length of time between refueling periods is mainly determined by the available reactivity remaining in the core. The utility would normally want to minimize refueling time and schedule the outage at times when required replacement power costs would be the lowest. The actual fuel movement activities take about ten days with additional time required for the component removal and replacement tasks for access to the core. When other maintenance activities are also included, a typical outage will be about two months in duration. The length and frequency of refueling outages affects the availability of the unit and the cost of producing electricity. Approximately one-third of the fuel assemblies are replaced at each refueling. The actual fuel load patterns are pre-determined as part of the reload core physics design and safety analyses to produce an acceptable core configuration. The type of fuel loading scheme must consider the requirements and constraints of the utility. The refueling shuffle itself can potentially be on critical path. A nominal BWR shuffle may contain as many as 1000 shuffle steps (steps that are required for the discharge of old fuel and to bring in the new fuel). An efficient core shuffle plan, particularly if the shuffle is on critical path will allow the plant to be brought on-line earlier with a proportionate reduction in outage cost.

II.1 Core Design Shuffling Considerations

During the refueling, it is necessary to remove any assemblies that would exceed their burn up limits during the upcoming cycle and replace them with new fuel. It is important to consider which locations the new assemblies will occupy and the impact that the new fuel reactivity will have on the power distribution in the core. These factors, reactivity and power distribution, are considered in the design of the new fuel and core placement patterns (reload core design). The core placement pattern is the predetermined final core configuration that the outage shuffle is attempting to achieve. The reload designer determines the desired/required locations for the fuel. The shuffle planner determines the desired/required sequence of crane and core component movement steps to achieve the core pattern.

Pressurized Water Reactors (PWRs) and Boiling Water Reactors (BWRs) both have fuel assemblies that must be shuffled for optimum performance. The BWR has more assemblies per core with each assembly being of smaller dimensions. A large BWR will have over 500 fuel assemblies while a typical PWR may have about 200 fuel assemblies. In the PWR, the burnable poison rods, thimble plugs, sources and control rods are inserted into guide tubes in the assemblies and must therefore be considered in the reload design and shuffle plan. In the BWR, the control rods are inserted between fuel assemblies and are not required to be shuffled during the fuel shuffle. Since control rod replacement in a BWR does require removal of the adjacent fuel assemblies, this operation does impact the fuel shuffle plan.

II.2 Shuffle Planning

Once the design of the reload core has been established, the planning for the shuffle can begin. The goal of a core shuffle planner is to determine an efficient sequence of crane and fuel bundle movements so as to move the fuel assemblies from their present positions (initial core configuration) to the new positions (final core configuration) required for the next cycle of operation in the minimum amount of time including such considerations as minimizing tool changes. There are situations where the complete core is off-loaded for refueling. For those reactors with inserts in every fuel assembly or when vessel or fuel inspections are required it may be more efficient to perform a full core off-load with the insert shuffle being performed in the spent fuel pool. A complete off-load may remove part of the shuffle from critical path and it also allows more flexibility in reactor maintenance and inspection activities.

For the in-core shuffles, since initially there are no empty locations in the core, the first step is to select certain assemblies for removal. These assemblies would consist of discharged fuel or fuel assemblies that may require out-of-core inspections. Once a location is opened by removing a fuel assembly, the replacement assembly, either a new fuel assembly or a fuel assembly to remain resident in the core for the next cycle, is moved to the empty location. This move then frees up another hole into which the designated fuel assembly would be moved. This chain of moves would end when the empty location is filled by the required assembly. Since there are only a limited number of fuel types, this process consists of many short "chains" of possible moves. Chains can be worked in serial or in parallel, resulting in a large number of possible moves. In many cases, more than one fuel position is opened in the core to allow more flexibility in the shuffle planning. This can achieve a more efficient plan at the expense of larger number of possible moves to be considered.

The shuffle planner must also consider inserts that the fuel assemblies contain. Inserts (control rods, burnable poison rods, neutron sources and thimble plugs) may often require discharge, replacement, or relocation to another assembly. The shuffle of these items may occur while the fuel is in the core or may be done outside the core. In the case where the complete core is off-loaded, optimizing the placement of the fuel assemblies during the off-load in storage racks can significantly reduce the time required for the insert shuffle. Therefore, the most important and difficult part of the planning is to determine the best location for the assemblies in the spent fuel pool such that the subsequent insert shuffle is efficient. The fuel assembly shuffle is handled simply by loading the fuel assemblies into their final locations in the core.

Plan Strategies

The following provides some insight into the strategies incorporated in the Planning System for Core Shuffles. Each strategy is designed to provide a minimum time for the shuffle based on user inputs of time durations for individual strategy steps. Constraints on fuel or control rod support and refueling mast orientations are included as user selectable options for use in the planner.

1. PWR IN-CORE SHUFFLE

The PWR in-core shuffle will perform the fuel and insert shuffle in the core area to the extent possible considering plant equipment. The system is able to handle new fuel, resident fuel and discharge fuel along with control assemblies, burnable poison assemblies, thimble plugs and source assemblies. Plant equipment used will be defined by the user and may include a main and auxiliary refueling machine, control element exchange machine, upenders and transfer machine, spent fuel handling machine, new fuel elevator and overhead crane(s).

The shuffle plan would be based on reducing total time and minimizing tool changes. A typical sequence would first perform an insert shuffle, then a fuel shuffle and finally a shuffle of all the remaining inserts. New fuel would be brought to the core and discharge fuel would be taken to the spent fuel pool during the shuffle process.

2. PWR SPENT FUEL POOL SHUFFLE

The PWR spent fuel pool shuffle will perform the insert shuffle in the spent fuel pool area. The system is able to handle new fuel, resident fuel and discharge fuel along with control assemblies, burnable poison assemblies, thimble plugs and source assemblies. Plant equipment used will be defined by the user and may include a main and auxiliary refueling machine, control element exchange machine, upenders and transfer machine, spent fuel handling machine, new fuel elevator overhead crane(s), and assembly and insert tools.

Optimizing the placement of the fuel assemblies into the spent fuel pool will reduce the time required for the insert shuffle. The placement of the fuel assemblies and the insert shuffle will be performed as a follow on to the algorithms developed by Joseph Naser, et al (Reference 2). In this scenario, all fuel is placed in spent fuel pool racks in an array that allows efficient crane movement and minimizes required tool changes during the insert shuffle. New fuel may or may not be required to participate in the insert shuffle depending on the insert previously loaded into the new fuel assembly. The system will also perform the insert shuffle on any user designed fuel assembly storage pattern.

Core reload will be performed by installed or user defined sequences. Installed reload sequences will consider temporary placement of assemblies containing secondary sources near source range detectors as a priority for the reload.

3. BWR IN-CORE SHUFFLE

The BWR in-core shuffle involves no inserts to be shuffled but must accommodate control rod drive and local power range monitor maintenance. The system will be able to handle new fuel, resident fuel and discharge fuel. Plant equipment may consist of a refueling machine, fuel preparation machine, new fuel elevator and overhead crane.

The user may manually specify the number of holes to open at the beginning of the shuffle or allow the computer to select the holes. Computer selection of the holes will be based upon maintenance requirements (inspections, control rod or drive maintenance and local power range monitor maintenance activities).

The system uses a simple k-infinity averaging scheme for checks against a user specified limit in designing the shuffle sequence. The system will have an interface for use by the user as input for a shutdown margin verification calculation.

Shuffle Planning Constraints

The method of planning employed is a knowledge-based system which attempts to minimize the overall time needed to execute a shuffle plan. The solution is bounded by various plant constraints, plan evaluation criteria, and plan strategies, including (but not limited to) the following:

Planning Constraints:

a. Accessibility of core and spent fuel pool locations by different cranes and lifting tools.
b. In-core assembly support constraints.
c. Spent fuel pool criticality constraints.
d. Presence of control element during the process of fuel movement (BWR).
e. Constraints on shut down margin during the process of shuffling or reloading the core (BWR).
f. Constraints on moving assemblies in a certain order (i.e., in BWR's assemblies are processed in groups of four in a given sequence)

One of the most important shuffle constraints particularly for BWR's is that adequate shutdown margin (SDM) be maintained during the refueling. Shutdown margin is defined as the amount the reactor is shutdown (subcritical) below the point at which the reactor will undergo a self-sustaining fission process. This ensures that the reactor is sufficiently subcritical so as to prevent the possibility of an inadvertent criticality accident. SDM is maintained in the PWR by adding sufficient boron to the reactor coolant. Since boron is not used in the BWR, a verification of the SDM at each step of the shuffle is required. This requirement may be satisfied by an analysis of the worst case configuration using a 3-dimensional, multi-group calculation analysis code or by using an alternate calculation for each step. Any alternate calculation should be benchmarked to the 3-dimensional code for the refueling under consideration. A typical approach to the alternate calculation would be to perform a 2-dimensional, single group eigenvalue calculation using assembly specific k-infinities generated from the 3-dimensional code.

III. CORE FUEL SHUFFLE PLANNING SYSTEM DESCRIPTION

III.1 Overview

The Core Shuffle Planning System is a PC based system with many features providing users with flexibility and a variety of planning capabilities. The shuffle planning system is capable of producing complete shuffle plans (fuel crane movement sequences) automatically given the initial and final core configurations. The shuffle planning system can automatically generate shuffle plans for BWR and PWR power plants. The desired requirements for the system, which was sponsored by EPRI, were defined in conjunction with a utility advisory group of more than 30 utilities. A set of general requirements was defined that met the utility groups representative needs. The modular design and flexible software architecture of the system allow it to be further tailored to a given utility's additional needs.

The shuffle planning system has the capability of interactively creating and/or modifying a shuffle plan as well as developing a complete plan automatically. Once a plan has been created, there is a facility for verifying the plan by interactively "walking through" the steps of the plan graphically on the computer screen and making changes as desired. This capability also allows for more accurate and faster evaluations of the plan for reviews and sign-offs as needed.

The shuffle planning system can produce the fuel handling sheets and core and spent fuel pool maps used by operators to perform shuffles during an outage. The system is very flexible in handling the wide variations in plant characteristics, equipment and constraints found at different sites. Some of the variations handled by the shuffle planning system include: user defined, arbitrary shaped Item Control Areas (i.e., any area which can contain nuclear material); any number of cranes in the core, spent fuel pool, and so on; user-definable insert types and tools for latching them; and arbitrary plant layouts. This is only a partial list of variations the system has been designed to handle.

The shuffle planning system has capabilities for monitoring the on-line execution of a shuffle during an outage. The on-line tracking ability allows control room personnel to keep track of floor area actions and keep an update on status, while maintaining a time history and log record of the job. In addition, it has many facilities for modifying shuffle plans or portions thereof due to problems encountered during the actual outage shuffle. These features are interactive and provide many aids for the automatic and semiautomatic replanning needed to deal with problems encountered in a quick and efficient manner.

The shuffle planning system has been designed to interface with existing fuel accountability systems through the use of standard format interface files. This allows easy definitions of the initial core and pool configurations as well as efficient means to supply the final configurations to the accountability system.

Finally, the shuffle planning system has an easy to learn and use user interface using multi-windowing, graphic, mouse-based interface technology. The user interface is intuitive with context-sensitive help available at all times.

III.2 Core Shuffle Planning Software Task Flow Description

Overview

This section describes the flow of tasks as the system is used to perform all of its functions. It provides a general overview of how a person would use the system to plan shuffles, perform on-line shuffle monitoring, and use the other features of the system. Although the following figures which represent system screens are black and white the actual screens are full color graphics.

Initial Set Up

For first time use, the user would start by selecting the System menu to define the characteristics of the power plant (see Figure 1). This includes picking the core model and defining the shapes and locations of the other ICA's (Item Control Areas). An Item Control Area is defined as any area in a plant which can contain nuclear material (e.g. core, spent fuel pool, new fuel storage racks, upender, inspection stand, and so on). ICA shape definition can be created graphically by moving ICA building blocks on the screen with the mouse to define the shape of an ICA. ICA's can have any arbitrary shape. Other set-up information includes plant equipment, type of shuffle desired, shuffle planning constraints, and so on. The power plant set-up information is saved in a file for later use and future shuffle plan development.

After the basic plant configurations have been defined, the user accesses the Set-up menu to load the initial core, spent fuel pool, new fuel storage and final core configurations in preparation for each shuffle.

Display Configurations

Once all power plant configurations have been loaded, the user can select the Display menu to display any desired ICA. This would probably include the core and/or spent fuel pool depending on which type of shuffle is being planned. Multiple ICA displays can be viewed at the same time (Figure 2).

ICA's can be displayed at two levels of detail. The full detail view displays an ICA with cells large enough to show assembly and insert serial numbers within each cell (Figure 3). This view allows all the details of traditional core maps to be seen on the screen. However, the amount of a complete core or spent fuel pool seen on the screen at one time is limited by the size of the screen. Large screens can be used to advantage to view more of the item control areas at one time.

The second level of viewing is a space saving micro view (Figure 4) with very small cells that can contain small black squares showing that a cell is occupied. When an occupied cell is pointed to with the mouse, the assembly and insert serial numbers are dynamically displayed in the message areas of the display. The micro view has the advantage that a whole core and much of a spent fuel pool can be displayed at the same time. In addition, each display window can be moved, resized and scrolled to view all portions of an ICA. Both views also have a color coding feature to point out the previous and current movement steps in an obvious manner.

Shuffle Planning

The shuffle planning module handles the automatic planning of shuffle sequences. It consists of several independent submodules used for planning different kinds of shuffles and for piecing together shuffle sequences. For instance, there are three different submodules for producing: PWR in-core shuffles, PWR off-load/insert shuffles, BWR in-core shuffles.

There are special modes for automating common fuel movement tasks. This includes, for example, moving a batch of new fuel from the new fuel storage racks to the spent fuel pool, moving assemblies one-by-one to an inspection site, and re-racking assemblies in the spent fuel pool. There is also a provision for entering steps interactively to handle arbitrary fuel movements. Complete shuffle plans are saved to files for later use.

The modules for automatically generating shuffle plans have the ability to start the planning process from an intermediate state of the shuffling process. This handles, for example, cases where the user has entered some initial moves manually and the shuffle system is intended to generate a plan from there, or where the system creates an initial plan, the user interactively inserts a step or sequence of steps and then the system finishes the plan. It is also useful for the situation where conditions change during the refueling requiring a significant modification of the remainder of the plan.

User Planning

The user enters the shuffle planning module from the main menu by choosing the "Shuffle" pulldown menu. At this point the system displays the values of all parameters that pertain to shuffle planning and asks the user if these values are acceptable. If not, the user is then advised to set these parameters in the set-up module. If the parameters are acceptable, then another menu of shuffle submodules is presented. These submodules are used to plan shuffle sequences.

In its simplest form, the user would pick one of the three main shuffle scenarios (e.g., PWR in-core shuffle, PWR off-load/insert shuffle, or BWR in-core shuffle), and the system would automatically generate a complete shuffle sequence. The internal shuffle sequence can then be added to, modified and/or saved in a file for later use.

In a more complicated case, the user may wish to piece together different shuffle sequences created using the available shuffle submodules. For instance, the user may use the interactive mode to enter some initial moves. The user could pick the PWR off-load/insert shuffle submodule to automatically generate the rest of the shuffle from there. Finally, the user might choose to insert an inspection sequence, using the inspection submodule, right after the core off-load portion of the overall shuffle sequence. All of these sub-sequences are appended/inserted together to form the complete shuffle sequence.

Multiple shuffle plans can be produced for comparison purposes and for "what if" purposes during planning.

Shuffle Plan Verification

Once a shuffle plan has been created, the user may want to visually "step through" the plan on the screen to verify the correctness and reasonableness of the plan. This can be done independently of whether the plan was generated automatically, entered interactively or a combination of both. The graphic verification module takes an arbitrary plan as input and animates the execution of the plan on the screen (Figure 5). The plan is checked automatically by the system for legality on a move-by-move basis. Checks such as the physical reasonableness of a step and potential constraint violations are performed. Additionally, this visual capability allows the user to evaluate the plan subjectively. This capability is also very beneficial after the plan has been completed for the formal verifications of the plan by reviewers other than the plan developer. The visual capability is much faster and more accurate than a manual verification done by moving magnets or paper representing the fuel assemblies and inserts.

Interactive Shuffle Planning and Modification

There are extensive facilities for interactive planning and modification of shuffle plans. These include operations at the sequence level where sequences can be created, deleted, concatenated, spliced and copied. Then there are operations at the individual step level for adding steps, deleting steps, modifying steps, searching for steps and so on. All operations use the same intuitive mouse-driven interface and menus, with on-line help capabilities.

On-line Outage Monitoring and Modification

During the outage, the On-line Monitoring module is used to track and monitor the execution of the shuffle plan. The desired shuffle sequence is recalled from its saved file, and the shuffle plan is presented step-by-step to the user. The user indicates to the system the start and completion of each step. The computer automatically stamps the time and date on the step for record keeping purposes. In addition to presenting the plan steps, the user is able to perform any needed changes to the shuffle sequences to handle problems that arise during the outage.

At any point during the shuffle process, the current state of the shuffle can be saved and restarted later. The usual shuffle process bookkeeping is also handled by this module (i.e., saving completed state, time and date and user sign-offs, change logs and so on). Upon completion of the execution of the plan, the results are available for reporting and for sending the information back to the accountability system.

Printing and Reports

The shuffle planning system is capable of producing a variety of reports and printed output. After a plan or plans have been generated, the Report menu is selected to print statistics about the total number of steps in the plan and the estimated time to execute the plan. The shuffle planning system prints, in a generic format, the final fuel handling data sheets used by operators during the shuffle.

At any time, the user can use the capabilities within the Reports menu to print the configurations of any of the ICA's. The initial, current (intermediate state) and final configurations can be printed. These maps would be printed for use during the on-line shuffle process.

Once the outage shuffle is completed, the Reports capability can be used to print final ICA configurations, the actual shuffle steps performed, and nuclear material movement histories.

System Requirements

The Shuffle Planning System is designed to run on 80286 based IBM^{TM} PC, PS/2 or compatible with at least 10 megabytes of extended memory and a 40 megabyte hard disk. An EGA graphics card with color monitor is also required. Preferred features include a VGA graphics card with monitor and a 80386 processor.

Additionally, a super VGA card with a 19 inch color monitor is useful. The 19 inch display is desirable for showing more of the power plant's components on the screen at one time, but is not necessary.

III.3 Benefits of AI Implementation

After interviewing several nuclear engineers at different utilities who plan shuffles, it was discovered that shuffle planning, as typically performed, is generally a procedural process where experience-based heuristics have already been incorporated into the procedure. The shuffle planning system described in this paper implements these procedural approaches where appropriate, and enhances them with AI techniques to make the system more flexible and able to handle all of the variations encountered in different power plants. In some cases, the same procedures as used by engineers were implemented but enhanced with AI techniques. In other cases, AI approaches were used instead of the procedural approaches used by engineers. These cases will be described in the next section.

The shuffling planning system has been developed in Common LISP using AI techniques. The use of LISP enhanced the productivity of the software development effort in addition to being used to implement the AI portions of the system. Common LISP contains features that are very useful for easily operating on groups of objects used by the shuffle planning system such as Item Control Areas, fuel assemblies, fuel assembly inserts, cranes, insert latching tools.

The Common LISP language in conjunction with the Gold Hill Windows^{TM} extension to Common LISP also made the development of the sophisticated user interface much easier to implement. The user interface was developed using Gold Hill Windows which is a high level interface to Microsoft Windows™, a multi-windowing, mouse-based environment (resembling the environment on the Macintosh™ computer). This resulted in an easy to learn and use system.

As mentioned earlier the shuffle planning system has been made more flexible through the use of AI techniques. The shuffle planning system is able to avoid making limiting assumptions about power plant characteristics and equipment used during a shuffle. The system is very flexible in handling the many variations among power plants. The user can specify the number and types of equipment available for performing shuffles including the ability to define new tools and fuel components. For instance, the user can specify the number and types of cranes located in the core and spent fuel pool and the use and coordination of the multiple cranes is handled by an intelligent scheduling module.

Use of AI Enhancements in the Shuffle Planning Modules

It was described earlier that the procedural approaches used by engineers in shuffle planning were, in some cases, enhanced with AI techniques and replaced by AI approaches in other cases. This section will describe in more detail the use of AI in the three shuffle planning modules discussed earlier (i.e. PWR in-core shuffles, PWR off-load/reload shuffles, and BWR in-core shuffles).

In all three modules, AI techniques are used to make the system more flexible in handling plant variations. One example of this is the coordination and use of multiple cranes in the core and spent fuel pool. Some utilities have more than one fuel movement crane in each of these areas. The shuffle planning system uses an agenda-based scheduling module to handle the use and coordination of different cranes. This is done by creating a description of each crane including: the location of the crane, the area(s) the crane can reach, the type of tasks the crane can perform, the time it takes to perform its tasks, whether or not the crane is currently available for use, and conflicts with the use of other cranes. The scheduler puts each crane on the agenda and maintains a simulated clock. The scheduler plans the use of the cranes based on the availability of each crane as they are simulated performing their tasks. This allows the system to flexibly use any number of cranes that a particular power plant may have in each area. Other plant variations like the types of fuel inserts, latching tools, and so on are also made more flexible using AI techniques and apply to all three shuffling modules.

In the case of PWR in-core shuffles, the shuffle planning system uses a fairly procedural approach similar to the way engineers plan shuffles. The PWR in-core shuffle planning procedure is enhanced by the AI techniques described above. The procedure is based on discharging a subset of the spent fuel bundles to create holes in the core, shuffling the remaining assemblies, and bringing in new fuel. At each point during the planning process, there are a set of candidate assemblies that can be moved into the available holes in the core. At each point the assembly which can be moved in the shortest time is picked. The time to move an assembly is based on avoiding changes of direction and distance calculations.

In the case of PWR off-load/reload shuffles, the procedural approach used by engineers was replaced by a more efficient AI based approach. AI techniques were used to determine the placement of assemblies in the spent fuel pool which minimizes the distance traveled moving each insert during the insert shuffle. Also, AI tree searching techniques were used to determine the optimal usage order of insert latching tools to minimize the change-out of different tools during the insert shuffle. These approaches are most relevant to plants which have several different types of fuel inserts. The resulting insert shuffle is more efficient than those usually produced by engineers.

BWR in-core shuffle planning involves a goal-directed subcomponent during the in-core shuffling of fuel assemblies. In addition to the general goal of shuffling the initial core configuration to the final core configuration, the BWR planning engineer must achieve the subgoals of opening up specific areas within the core. This may be the case when control rod drives or power range monitors need to be serviced; the assemblies surrounding them must all be removed. Another example would include performing an inspection of a region of the core vessel. The removal of these assemblies is a subgoal that must be achieved during the overall process of core shuffling. The shuffle planning system uses an AI based approach of subgoal planning to flexibly achieve these subgoals.

IV. CONCLUSION

The paper has described a new and comprehensive core shuffle planning system that incorporates traditional shuffle planning procedural approaches with some AI software techniques to provide a more general and flexible enhanced capability. This capability allows planners to handle a variety of plant configurations, constraints and equipment that may be encountered at any given time or plant site. In addition to the planning functionality, the system provides for on-line monitoring to facilitate tracking and maintaining a record of the fuel movement portion of the outage. The shuffle verification module provides animated playback of shuffle plans for verification reviews. An interactive mode allows creating and/or modifying a shuffle plan. This mode allows "what if" planning sessions. Also on-line modifications to a shuffle plan can be made during an outage should problems occur with a given move (e.g., bent fuel bundle) allowing new moves and a modified plan to be generated quickly and accurately. The animation and interactive modes could also be used for training purposes allowing for dry-runs of fuel shuffle sequences.

The system provides hardcopy reports, shutdown margin calculation constraints and interfaces to separate criticality calculations and nuclear fuel accountability systems.

The benefits of the total capabilities provided in the planning tool include: faster development of plans; more efficient plans; automated checking and verification of plans; faster modification of plans (particularly during outages, if necessary); potential for reduction of refuel outage time, on-line tracking and record keeping during the outage. Also the system can be used in the interactive and animation modes as a training tool for utility engineers and outage personnel.

ACKNOWLEDGEMENTS

The authors acknowledge the technical contributions to the development of the Nuclear Fuel Shuffle Planning System by many colleagues within C-E's Nuclear Power Businesses and the Utility community. Particular thanks to the software design team: Matthew Allen, Sin-Kie Poon, Peter Rzasa and Lynette Smith. Additional thanks to Mohammed Elmeghrabi, Crawford Fountain, Wayne Johnson, Bill Young and Ed Ruzkauskus for their technical familiarity with the shuffle process and their valuable input. Further thanks to Gary Marsh of Florida Power & Light, Tom Brookmire of Virginia Power, Ron Furia and Howard Crawford of General Public Utilities, Robert Borchert of Northeast Utilities and the other utility representatives who have participated in the design and have supported the project.

REFERENCES

1. A Fuel Insert Shuffle Planner System, Joseph Naser, Robert Colley, John Gaiser, Thomas Brookmire, and Steven Engle, EPRI Seminar Expert Systems Applications in Power Plants, Boston Ma. May 1987.

2. Fuel Insert Shuffler: A Case Study of Expert System Development, Joseph Naser, Robert Colley, John Gaiser, Thomas Brookmire, and Steven Engle, ANS Topical Meeting on AI and other Innovative Computer Applications, Snowbird Utah, Aug 31 - Sept. 2, 1987.

3. The Development of CLEO; A Logic-Based Refueling Assistant for the FFTF, D. E. Smith, and L. F. Kocker, Hanford Engineering Development Laboratory.

4. CENMAS™, Combustion Engineering Nuclear Material Accountability System, Internal C.E. report.

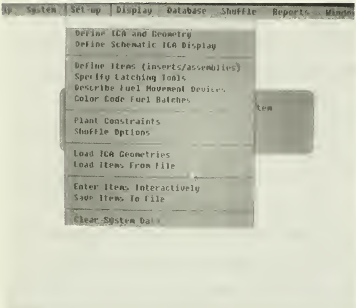


Fig. 1: Top Level Menu and Selected Setup Submenu

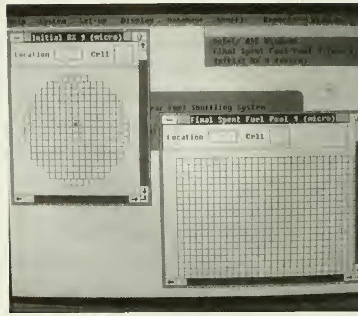


Fig. 2: Multiple ICA Displays (Cells Empty)

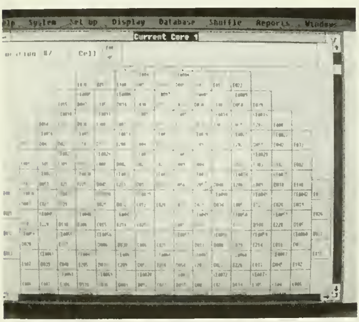


Fig. 3: Display Menu, Full Size Core

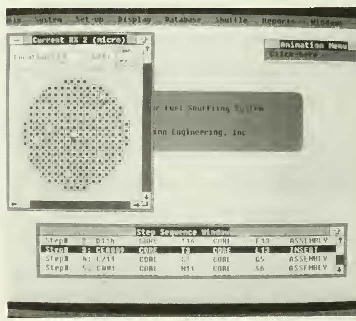


Fig. 4: Microview of Current Core and Sequence Step

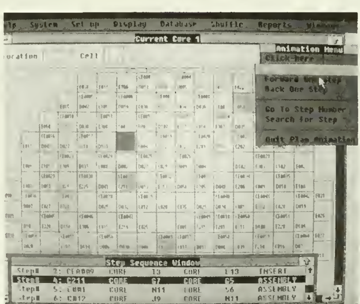


Fig. 5: Shuffle Plan Animation Using Full Display

Fluid Component Review for Age-Related Degradation

STAN SMITH
Yankee Atomic Electric Company
580 Main Street
Bolton, Massachusetts 01740, USA

ABSTRACT

Understanding, identifying and managing the different ways in which fluid system components can degrade when exposed to their environments is one of the more substantial elements of developing a technical basis for license extension, or PLEX. However, performing detailed evaluations of the tens of thousands of components within a power plant to identify how the component's environment will cause the component to age would be a very time consuming and tedious task, if done manually. To automate these decision processes, Yankee Atomic Electric Company (Yankee Atomic) developed an expert system which was used to review the fluid system components at the Yankee plant. This tool was used in 1988 to evaluate selected components (780) in 30 different fluid systems to determine the scope of age-related degradation and provide direction for future work associated with PLEX. The expert system is called CoDAT (Component Degradation Assessment Tool), and based on the 1988 evaluation results it is presently being updated to perform a more detailed evaluation of all Yankee plant fluid components. The results of this more detailed review will be published in the EPRI/DOE sponsored Lead PWR Plant Life Extension Project in January 1990.

INTRODUCTION

Managing fluid component age-related degradation requires a thorough understanding of all the ways a component can degrade due to its environment. Once this knowledge is obtained, utilities will be able to identify where in the plant the potential for fluid component degradation exists and take the necessary actions to monitor the progression of the degradation.

For the past two years, Yankee Atomic has been gathering information from other operating plants, as well as our own, and industry reports related to age degradation of fluid components. As a result of this research, we have obtained an excellent understanding of fluid component degradation. The knowledge gained during this process has been represented in the form of "logic diagrams", from which simplified rules were developed and used in the development of the expert system.

The name of the expert system is Component Degradation Assessment Tool, or CoDAT. CoDAT can operate in two different modes. In the automatic mode, it accesses several data bases that store the special parameters necessary to predict age-related degradation. Because all the information required to evaluate the component for degradation is in the data bases, the entire evaluation process is automatic. In the second mode of operation, or user mode, the user is required to enter information as the expert system determines the need for the information.

EXPERT SYSTEM APPLICATION DESCRIPTION

PLEX THEORY

Their are over 100 operating commercial nuclear plants in the U.S. today. Several of these power plants have been operating for over 20 years and are approaching the end of their licensed operating period. For these older utilities, plans for construction of replacement power must soon be addressed. One way to help meet the energy needs of the future and defer the cost of new construction is the Plant Life Extension option, or PLEX. PLEX offers utilities the choice of extending their operating license provided they can effectively manage degradation of plant systems and components.

FLUID COMPONENT ANALYSIS

The tools required to show that degradation of fluid systems components is managed effectively are a good understanding of the ways in which the components can degrade and a uniform method for determining where this degradation may occur due to the component's operating environment. For the fluid systems at Yankee, we identified 18 groups (28 specific) of degradation mechanisms that could cause fluid components to degrade. The 28 degradation mechanisms do not include such initiators as improper welding techniques, torquing, cleaning, maintenance, etc.

DEGRADATION MECHANISMS

The 28 degradation mechanisms that could affect the fluid systems at Yankee are listed in Table 1 (these degradation mechanisms are grouped under 18 major headings). These mechanisms were selected from an EPRI Report titled, Component Life Estimation:LWR Structural Materials Degradation Mechanisms, NP-5461 and from the Yankee plant operating experiences. Not all of the mechanisms listed in the EPRI report were applicable to the Yankee operating environment. For instance, creep is a time dependent strain which occurs under stress. However, research and experience indicate that certain conditions must be met before this strain will occur. One condition which must be present is a component operating temperature greater than 1100 F (for carbon steels). For a typical pressurized water reactor (PWR), which operates at about 600 F (like Yankee), creep would not be considered a mechanism which could cause degradation of fluid components.

Of the 18 degradation mechanism groups applicable to Yankee, we felt that only 14 of these groups (21 specific degradation mechanism) could be evaluated using an automated reasoning tool like an expert system. For the seven remaining mechanisms, we determined that they could be more efficiently addressed by reviewing the present component surveillance activities, using already developed commercial software, or performing system walk downs. These 7 mechanisms are marked with an "*" in Table 1.

CAPTURING KNOWLEDGE

INFORMATION SOURCES

After determining the degradation mechanisms which could be applicable to the Yankee environment, a search was performed to gain further knowledge of the 28 degradation mechanisms. The search produced a list of information sources which were found to be helpful in predicting degradation of a fluid component (These references are listed in the REFERENCES section of this paper). Many information sources, in addition to those discussed above, were also reviewed. However, they were not included in this list because they were either lacking in detail or they discussed a specific problem, the results of which, could not be easily generalized.

CONTROLLING PARAMETERS

During the degradation mechanism review process, Yankee identified some special parameters that were useful in predicting a component's susceptibility to degradation. We called these parameters Controlling Parameters, because they control whether or not a degradation mechanism could potentially exist, depending upon its value. For the degradation mechanisms applicable to the Yankee plant, we found that all of the controlling parameters could be classified into one of two categories. These two categories are identified as,

Component Material Characteristics, and

Operating Environments.

Based upon our review of the mechanisms applicable to Yankee, forty one controlling parameters were determined to be effective in predicting fluid component degradation. A list of these controlling parameters is shown in Table 2.

LOGIC DIAGRAM REPRESENTATION

Knowing that we would probably build an expert system, representation of the knowledge obtained from our research became important, because the method in which we documented the knowledge must be easily converted to the "if-then" format used by many expert system shells. Examples of these logic diagrams are shown in Figures 1 and 2. These diagrams identify the acceptable path(s) that a system engineer may use to determine when a fluid component may degrade due to its environment. The diagrams also identify the controlling parameters, the acceptable values for these parameters, and the information required to reach a decision.

Fourteen degradation mechanism logic diagrams (one for each major group evaluated by CoDAT, shown in Table 1) were developed to perform the screening evaluation at Yankee. An independent review of the technical bases supporting the logic diagrams was performed by an outside party.

EXPERTS SYSTEM DESCRIPTION

PURPOSE OF SYSTEM

The Component Degradation Assessment Tool, or CoDAT, was originally developed to aid in the determination of fluid system component degradation, and by doing so, aid in the scheduling of future work related to PLEX. CoDAT achieved this goal by performing a screening of selected components from 30 different systems (780 components total). Based upon the screening results, CoDAT is being revised to permit an analysis of all plant fluid components determined to be safety related or otherwise important to plant operation.

Since the evaluation of fluid components even with the aid of an expert system is complicated, CoDAT was designed to be used only by engineers, operators or maintenance personnel knowledgeable in fluid system operating conditions and fluid component material characteristics. It can be operated in two different ways or modes. In the first mode, CoDAT accesses information stored in data bases and uses this information to evaluate the plant's fluid components for degradation due to aging. This mode is referred to as the "automatic" mode.

One problem which we encountered while using the automatic mode, was incorrect or misspelled data in the data bases. Since CoDAT could not recognize this data, the results were not what we expected. We solved this problem by placing controls on the data going into the data base and checking it prior to use in CoDAT. Since checking data for thousands of fluid components can be time consuming, we decided to design a subprogram for CoDAT that would perform the job. This subprogram checks each piece of data important to the degradation evaluations against a list of acceptable values for that data type. The subprogram was designed to aid the persons supplying and inputing the data by identifying the specific record(s) and data field(s) which were incorrect. The data check program is performed prior to CoDAT being used in the automatic mode. In addition, included in the CoDAT knowledge base are rule conclusions which also warn the user that an unrecognizable process fluid type or material classification exists and that specific rules have not been developed to evaluate this specific case (this feature was initially added as a debugging aid, however, it was left in the rules because it identifies when and where additional development is required).

The second mode of operation is called the "user" mode. In this mode, the user is asked to supply the information requested by the expert system. The advantage of this operating mode is that only the information required to provide a result are gathered, where as, in the automatic mode of operation some of the information gathered may never be used by CoDAT. In the user mode of operation, data entry errors are eliminated because in most cases the user selects the appropriate answer from a menu generated for each question asked. Since numeric answers are not conducive to the development of a menu, the appropriate range for the numeric value is monitored by CoDAT. As an example, when CoDAT requests that the user enter a pH value for the process fluid, it will not accept a value outside of 0-14. If the user tries to enter 15 as a pH value, CoDAT informs the user that the acceptable range is 0-14 and requests the value for pH be reentered.

EXPERT SYSTEM SHELL DESCRIPTION

CoDAT was initially developed on a commercial expert system shell. The shell was purchased for approximately $99. Some specific attributes of the shell are identified below:

- + Operates on an IBM PC, XT, AT and most clones with 256K or more of RAM memory, one disk drive and DOS version 2.0 or higher

- + The ability to exchange data with VP-Info or dBASE files (up to III+), VP-Planner or Lotus 123 worksheet files, and ASCII text

- + An inference engine that uses backward and forward chaining for problem solving

- + Confidence factors that let you account for uncertain information in a knowledge base

- + Simple English rule construction

- + The ability to explain its actions during a consultation

- + Knowledge base size limited to 32K of ram

- + Knowledge base "chaining" which lets you create knowledge bases that would otherwise be too large to fit into memory
- + A built in text editor

- + Ability to access up to 6 data bases at any one time

Because of limits in knowledge base size and some difficulties related to accessing specific information in data bases, Yankee Atomic is presently converting the rules contained in CoDAT to another commercial expert system shell better suited for our application.

Rule Format

The rule format utilized by the system shell is a simple IF-THEN format, structured as shown in Figure 3. As shown in this figure, up to 20 conditions can be listed under the premise (if statement) of a rule. Any number of conclusions and/or clauses can follow the conclusion (then statement) of the rule.

Else and because statements can also be used (if desired) in the rule format. The else statement follows the conclusion of the rule and is only accessed if the rule does not pass. The because statement allows the programer to provide a message to the user explaining how the conclusion was reached.

There are approximately 350 rules in CoDAT. Three hundred and seventeen rules determine whether a component may experience degradation and the remainder are used to check the data base for data entry errors and control program direction. The 317 rules which determine if degradation may occur are sectioned into the 14 major degradation mechanism headings and represent the logic diagrams.

DATA BASE FORMAT

When CoDAT was first used in 1988, it accessed one large data base, which contained both the input data required to determine if any of the 21 degradation mechanisms would cause fluid component degradation, and the output data, which contained the results of the evaluations. The data base had approximately one hundred fields. Presently, CoDAT accesses 11 data bases from which input data is retrieved and 1 data base which receives the results. The relation between the data bases and CoDAT are shown in Figure 4.

EVALUATION RESULTS

The results of the preliminary evaluation performed in 1988 indicate that 93 percent of the potential degradation concerns, for the 780 components, have been eliminated. The results of the 10,920 (780 components x 14 major groups of degradation mechanisms) evaluations have been documented using coding which refers the reviewer back to the rule which was used to reach the evaluation conclusion. The remaining seven percent represent areas where more detailed evaluations are required to determine the true impact to PLEX. These areas are being evaluated to ensure the existing preventative maintenance, surveillance and/or inspection practices performed at Yankee can effectively manage the potential degradation mechanisms. Where the present practices are not completely effective, the results obtained from the screening evaluation will be used to define more effective surveillance and preventative maintenance practices.

Since the preliminary evaluation at Yankee looked at all systems and many different components within each system (not just at systems or components which were suspected of a particular degradation mechanism), some of the results were unexpected. For instance, one generally accepted industry guideline (NRCB 87-01, Thinning Of Pipe Walls In Nuclear Power Plants) used to limit the scope of evaluations required to determine if erosion/corrosion (E/C) can exist is based on system operating temperatures being between 190 - 500 F. Where temperatures outside this range are considered to produce negligible wall thinning. Systems which operate above the 500 F may not be reviewed for E/C, even though all other conditions required for E/C are met. CoDAT's rules for E/C did not include the upper temperature of 500 F because we felt any wall thinning of a carbon steel, high energy system was unacceptable. As a result, CoDAT identified E/C as a potential degradation mechanism for the Steam Generator Blowdown System. During the last refueling outage in November of 1988, CoDAT's results were confirmed when a leak occurred during a system hydrostatic test of the blowdown system. Further evaluation for the extent of wall thinning indicated that E/C and possibly two phase erosion were concerns for the Yankee blowdown system. Appropriate steps are being taken to monitor the progression of this degradation.

CONCLUSION

The utilities industry has learned a great deal about the safe operation of its power plants in the last hundred years. However, much of the time, the information is not always effectively disseminated and the experts end up being the only people who really know what's going on. Since the experts are few in number, it makes sense to capture their knowledge using an expert system tool such as CoDAT.

CoDAT has demonstrated its value in identifying the areas of the plant where more detailed attention to fluid system degradation is warranted. Of equal importance, it provides a formal and expedient process of documenting the areas of no concern.

FIGURE 1
THERMAL EMBRITTLEMENT

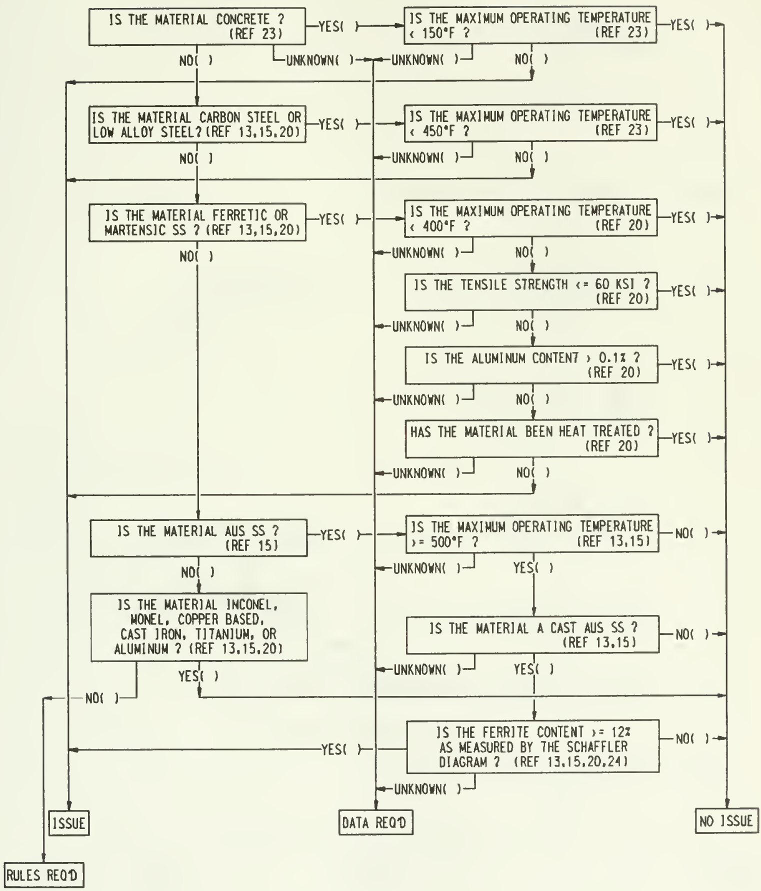


FIGURE 2
SINGLE PHASE FLOW EROSION/CORROSION

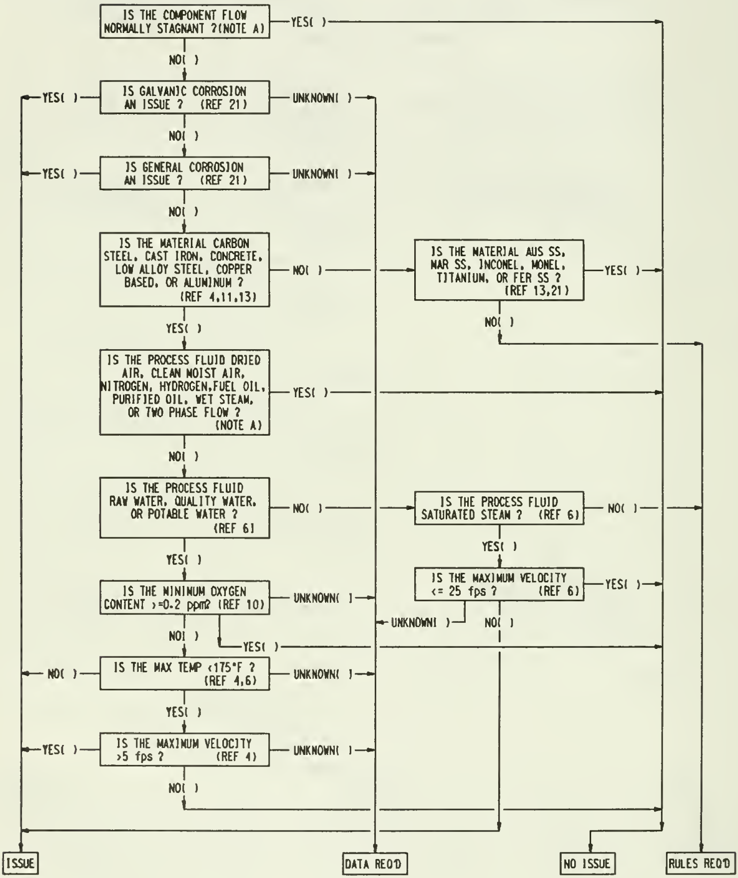


FIGURE 3
TYPICAL RULE FORMAT

RULE NAME —— RULE IGSCC #24

RULE PREMISES (UP TO 20) ——

IF MATLCLASS = AUS SS AND
MATLTYPE = 304L OR
MATLTYPE = 304NG OR
MATLTYPE = 316L OR
MATLTYPE = 316NG OR
MATLTYPE = 347NG OR
MATLTYPE = CF3 OR
MATLTYPE = CF3M

RULE CONCLUSION —— THEN IGSCC = NO PROBLEM

OPTIONAL —— ELSE IGSCC = MAY BE PROBLEM;
(IF RULE DOES NOT PASS, ELSE STATEMENT IS EXECUTED)

LEFT OF " = " ARE VARIABLES (CONTROLLING PARAMETERS)

RIGHT OF " = " ARE VALUES FOR THE VARIABLES

FIGURE 4
CoDAT RELATIONSHIP WITH DATA BASES

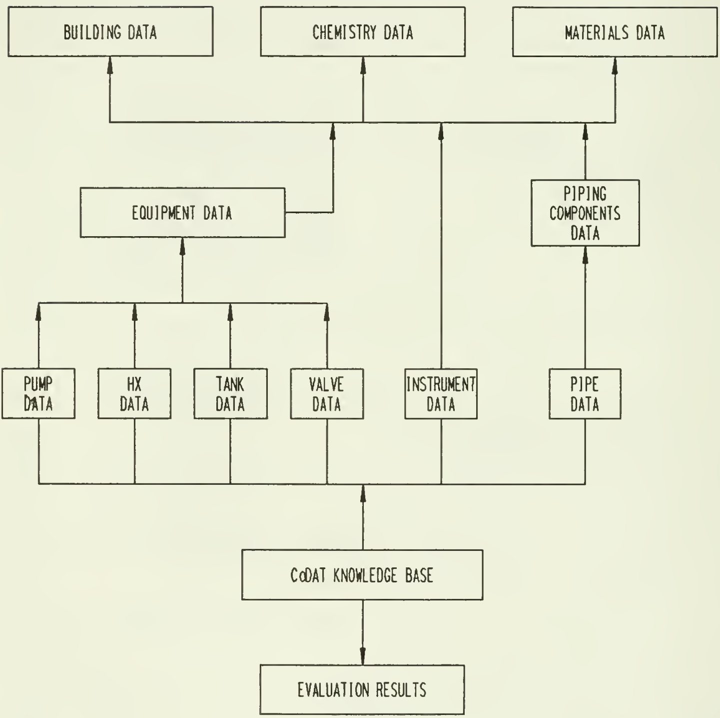


TABLE 1

<u>Fluid Component Degradation Mechanisms Considered For PLEX</u>

- General or Uniform Corrosion
- Erosion/Corrosion
- Two Phase Erosion
- Microbiologically Influenced Corrosion
- Intergranular Stress Corrosion Cracking
- Transgranular Stress Corrosion Cracking
- Irradiation Assisted Stress Corrosion Cracking
- Intergranular Attack
  - Knifeline Attack
  - Weld Decay
- Crevice/Pitting Corrosion
- * Thermal Fatigue
- Thermal Embrittlement
  - 885 F Embrittlement
  - Strain Age Embrittlement
  - Blue Brittleness
  - Temper Embrittlement
  - Quench Age Embrittlement
- Irradiation Embrittlement
- Hydrogen Embrittlement
- Selective Leaching
  - Dezincification
  - Graphitization
- Galvanic Corrosion
- * Wear
  - Galling
  - Abrasion
  - Fretting
- * Mechanical Fatigue
  - Cyclic Loading
  - Vibration (Rotational)
  - Vibration (Flow Induced)
- * Lubrication Breakdown

* Degradation mechanisms <u>not</u> presently evaluated by CoDAT
\+ Analyzed by other, existing programs

TABLE 2
List Of Fluid Component Controlling Parameters

Operating Environment Parameters

Process Fluid Type	Chemicals Added To System
External Surface Environment	Cathodic Protection Used
System Treated For MIC	Fluid Chloride Content
Fluid pH Value	Fluid Fluoride Content
Fluid Conductivity	Fluid Oxygen Content
Potential For Impurity Concentration	Fluid Chromate Content
Fluid Boron Content	Operating Pressure
Saturation Pressure	Fluid Velocity
Maximum Temperature	Minimum Temperature
Lifetime Neutron Exposure	Lifetime Gamma Exposure
Internal Surface Coatings Used	System Operating Mode

Material Characteristic Parameters

General Classification	Code Description And Type
Welding Used	Special Material Treatments
Material Copper Content	Material Zinc Content
Material Aluminum Content	Material Chromium Content
Material Carbon Content	Equivalent Chromium Content
Material Molybdenum Content	Material Hardness
Equivalent Nickel Content	Material Ferrite Content
Galvanic Potential Rating	Adjacent Material Classification
Material Yield Strength	Material Tensile Strength

ACKNOWLEDGMENTS

Developing an expert system requires the involvement of a team. I would therefore, like to thank all of the Yankee personnel in the PLEX group for their ideas, especially project management, without whose support and trust the system would never have been developed. I would also like thank Eric Biemiller for his expertise and encouragement, and also Donna McClellan and Susan McConaty of Stone & Webster, who performed a lion's share of the work required to develop the basis for fluid component degradation and get the expert system up and running.

REFERENCES

1. Yankee Specific Reference

2. Engineering Materials And Their Applications by Richard A. Flinn and Paul K. Trojan, Copyright 1975 by Houghton Mifflin Company.

3. Nuclear Steam Supply Systems Chemistry Manual, CENPD-28, Rev. 2, Combustion Engineering Power Systems.

4. Erosion-Corrosion Experiments and Calculation Model by W. Kastner, Kraftwerk Union AG, presented at an EPRI sponsored Erosion/Corrosion Program, April 1987.

5. Manual for Determining the Remaining Strength of Corroded Pipelines, a supplement to ANSI/ASME B31 Code for Pressure Piping, ANSI/ASME B31G, 1984.

6. Yankee Specific Reference

7. Yankee Specific Reference

8. NUREG-1061, Volume 1.

9. Metallurgy of Welding, by J. F. Lancaster, 1980.

10. Corrosion-Related Failures in Feedwater Heaters, B. C. Syrett, EPRI Report CS-3184, July 1983.

11. Erosion/Corrosion in Nuclear Plant Steam Piping: Causes and Inspection Program Guidelines, N. S. Hirota, EPRI Report NP-3944, April 1985.

12. A Study of Microbiologically-Influenced Corrosion in Nuclear Power Plants and a Practical Guide for Countermeasures, D. Cubicciotti, EPRI Report NP-4582, May 1986.

13. Component Life Estimation: LWR Structural Materials Degradation Mechanisms, M. E. Lapides, EPRI Report NP-5461, Sept. 1987.

14. Microbial Corrosion in Fossil-Fired Power Plants - A Study of Microbiologically-Influenced Corrosion and a Practical Guide for Its Treatment

and Prevention, J. A. Bartz, EPRI Report CS-5495, Nov. 1987.

15. Environmental Effects on Components: Commentary for ASME Section III, S. W. Tagart, Jr., EPRI Report NP-5775, April 1988.

16. Proceedings: Workshop on Initiation of Stress Corrosion Cracking Under LWR Conditions, D. Cubicciotti, EPRI Report NP-5828, May 1988.

17. Corrosion Control in the Chemical Process Industries, by C. P. Dillon, Copyright 1986 by McGraw-Hill, Inc.

18. Corrosion Data Survey, National Association of Corrosion Engineers, Copyright 1974.

19. Corrosion Engineering by Mars G. Fontana, Copyright 1986 by McGraw-Hill, Inc.

20. Metals Handbook Desk Edition, by the American Society For Metals, Copyright 1985.

21. Galvanic Corrosion - Electrochemical Theory of Galvanic Corrosion, by Harvey P. Hack, editor, Copyright by American Society of Testing And Materials, 1988.

22. 1987 EPRI Workshop on Secondary - Side Intergranular Corrosion Mechanisms: Proceedings, J.P.N. Paine, EPRI Report NP-5971, Vol I and II, April 1987.

23. NUREG/CR-4652, Concrete Component Aging And Its Significance Relative to Life Extension of Nuclear Power Plants.

24. LWR Experience With Centrifugally Cast Stainless Steel Pipe, S.N. Liu, EPRI Report NP-4996-LD, Dec. 1986.

25. Stress Corrosion Cracking of Alloys 600 and 690 in All -Volatile - Treated Water at Elevated Temperatures, by C.E. Shoemaker, EPRI Report NP-5761SP, May 1988.

26. 10CFR50 Appendix G, Fracture Toughness, and Appendix H, Reactor Vessel Material Surveillance Program Requirements.

27. Regulatory Guide 1.99, Rev. 2, Radiation Embrittlement of Reactor Vessel Material.

28. VP-EXPERT Rule-Based Expert System Development Tool, by Brian Sawyer, Published by Paperback Software International 1987.

29. ASME Section III, Paragraph 2110c, 1983

PLEXSYS: An Expert System Development Tool for Electric Power Industry—Application and Evaluation

HIROSHI SAKAMOTO
Nippon Atomic Industry Group
4-1, Ukishima-cho, Kawasaki-ku
Kawasaki City, 210 Japan

MAOMI MAKINO and KIYOSHI TAKASAKA
Toshiba Corporation
8, Shinsugita-cho, Isogo-ku
Yokohama City, 235 Japan

DAVID G. CAIN and BILL K. H. SUN
Electric Power Research Institute
3412 Hillview Avenue
Palo Alto, California 94304, USA

ABSTRACT

PLEXSYS is an AI tool customized for use in electric power industry developed by Electric Power Research Institute (EPRI). Under cooperative agreement with EPRI, Toshiba Corp. participated in the project since 1986. The role of Toshiba is to; (a) support developing technical specifications reflecting experiences as nuclear power plant manufacturer, (b) evaluate capabilities of PLEXSYS through application to various typical engineering problems. The former goal have been accomplished by end of 1987 and research activities on the latter goal is currently under way. Two types of expert systems, Design Support Expert System and Diagnosis Support Expert System, have been developed by Toshiba for evaluation of PLEXSYS. Technical features of these systems and evaluation results on PLEXSYS are described in the paper.

INTRODUCTION

In electric power industry, demands for safety, reliability and economics are increasing year by year. These demands are particularly strong for nuclear power generation stations and many efforts to enhance reliability and efficiency of plants are taking place. One of these efforts are application of state of the art computers and digital information processing technologies in such fields as instrumentation, control, monitoring, communication, data acquisition, data base and others. Such systems take advantage of large mass of information using their enormous computing powers. However, since use of fully automated systems are still limited in nuclear power plants, engineers and operators of nuclear power stations are constantly exposed to quantitatively and qualitatively massive information.

PLEXSYS : An Expert System Development Tool for Electric Power Industry - APPLICATION & EVALUATION -

To decrease human burden on information processing, attempts to apply computers for more advanced purposes are coming to reality with help of artificial intelligence (AI) technology. Many such systems, often referred to as expert systems (ES), have been developed and some reaching practical level. Various AI method to transfer human knowledge into computers have been tested through prototype developments and turned out number of different approaches are possible to reach the goal. Yet to push technologies from laboratory into actual engineering fields standarization is an important factor for many reasons such as software productivity, training, maintenance, integration, technology transfer and so on.

Nuclear Power Division in Electric Power Research Institute (EPRI) initiated a research project to develop an expert system building tool named **PLEXSYS** (**PL**ant **EX**pert **SYS**tem) in 1985. (1) Under cooperative agreement with EPRI, Toshiba Corporation supported development of PLEXSYS since 1986. After completion of first phase on development of basic functions and technical specifications for future improvements in the end of 1987, Toshiba and EPRI entered second phase on evaluation of PLEXSYS through development of practical application systems. (2) Following part of this paper will summarize basic capabilities of PLEXSYS, describe features of application systems developed by Toshiba and conclude with the evaluation results derived from the application system development.

GENERAL FEATURES OF PLEXSYS

PLEXSYS is a software which provides a computer environment or platform for developing various types of expert systems. The project was originally initiated with intention to support engineers in electric power industry especially those working for nuclear power plants and PLEXSYS is designed to provide functions customized to support problem solvings in this particular field. Such type of AI software, a tool kit customized for use in certain domain, is often called a "domain shell" and PLEXSYS may be called "plant engineering domain shell".

Ideas of PLEXSYS is based on following simple observations.

(a) In electric power industry, engineers always pull out design drawings to solve problems and spend long time thinking on the drawings.

(b) There are many types of design drawings for power plants but any type of design drawings strictly follow their drawing principles.

(c) To read and solve problems, plant engineers make use of drawing principles, common sense and heuristics based on experience.

These observations suggest that design drawings play important role for problem solvings in electric power industry and a software platform with capabilities to represent information described on drawings and to use such information will be of great help for developing advanced expert systems. Basic paradigm dominating characteristic capabilities of PLEXSYS is called "Model Based Reasoning", a concept in AI often used in contrast with "Rule Based Reasoning" and in a word PLEXSYS is a software tool for building model based systems.

In rule based systems knowledge for solving problems is represented as rules best known in "If A then B" from, whereas in model based systems knowledge is represented as domain models. PLEXSYS models are characterized with following features.

(a) Models are simplified but general description of problem domain.
(b) Models consist of component objects with attributes and relations.
(c) Models have graphical representation equivalent to original drawings and also consistent with internal expression.

Model representation function of PLEXSYS (called **ModelEditor** modules) allow users to create models with simple graphical operations leaving the complicated internal data handling tasks to the system.

PLEXSYS models represent knowledge in form of network suited to express piping diagrams and electrical wirings. Since this knowledge representation is totally different from that of rules, reasoning mechanism to use such information is also necessary. Model based reasoning function of PLEXSYS (called **NetworkInspector** modules) provide capabilities to support solving problem directly from models without converting them to rules. Model based reasoning capability is unique and powerful characteristics of PLEXSYS suited for performing tasks combined with logical search among the model structure. Original PLEXSYS NetworkInspector without any modifications provides functions to read schematics like a novice engineer and more intelligent capabilities can be added through application developments. Ways to add new capabilities are either write additional piece of program into the NetworkInspector module or to make use of rules.

Although model based reasoning is the basic paradigm of PLEXSYS, it does not mean that model based reasoning is considered superior to rule based reasoning. Rules are powerful for representing heuristics or jumping over complicated logic and capabilities to combine models and rules are desired for developing practical expert systems. PLEXSYS does not have rule based reasoning function of its own, however it is built on top of general purpose AI tool **KEE** (**K**nowledge **E**ngineering **E**nvironment: commercial product of IntelliCorp) and can use full power of KEE including its reasoning mechanism. (Figure 1)

APPLICATION SYSTEMS

To evaluate the existing capabilities of PLEXSYS and also to pick up necessary improvements two application systems have been developed. One is an expert system for supporting system designs and/or design reviews, another is an expert system for supporting diagnosis of electrical devices in plant control systems. Features of these systems are described in this chapter. (Figure 2)

a.Design Support Expert System

Various types of design drawings are used in power generation stations and whenever any modification is required, plant engineers have to go through sheets

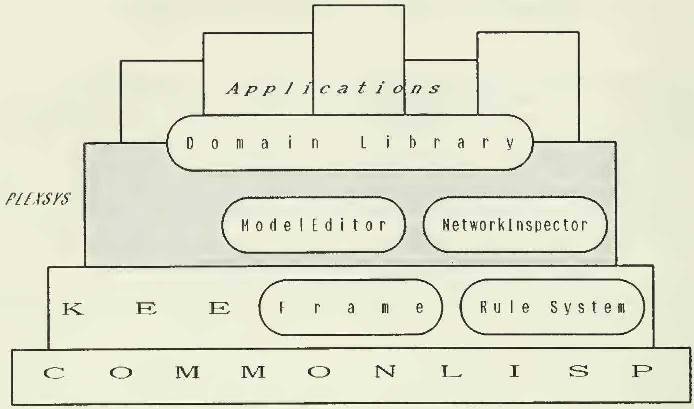


Figure 1. Software Structure of PLEXSYS

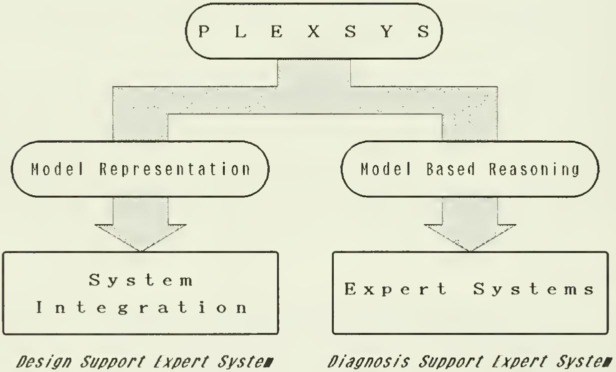


Figure 2. PLEXSYS Applications

of drawings for both finding out necessary changes and reviews. Especially in complicated systems like nuclear power plants even a slight modification may affect entire system functionality and careful evaluation on various types of design documents are necessary. CAD systems are being used for generating design documents recently, but most of these are advanced drafting systems and also can handle single type of drawings at a time. As a result, most of the work for design changes and their reviews are done by hand. These are time consuming works but important for maintaining reliability and safety of power plants. Expert system that can search through different types of design drawings and collect necessary information is expected to be a great help for engineers in making design changes and reviews.

The generic model representation capability and model based reasoning capability of PLEXSYS is suitable for such type of problem and a design support expert system using PLEXSYS was developed. Making use of flexible model representation capability of PLEXSYS, this system can handle information of various design documents on a single computer environment, such as **P&ID** (Piping and Instrumentation Diagram, Figure 3), **IBD** (Interlock Block Diagram, Figure 4) and more. The original capability of PLEXSYS provides functions to logically seek through these models and collect information under given conditions. In addition to these basic functions several other functions such as logical simulations, simple design calculations are added to support actual design works. The system was developed on AS workstation (alias of SUN workstation in Japanese market commercialized through Toshiba) and BWR plant High Pressure Core Spray (HPCS) system was selected as a test case.

Current design support system is built with more emphasis on reduction of human engineers than on automation. As a result design support functions of the system is initially developed to cover as wide variety of work as possible instead of going deep into each tasks. In this sense, current system is still in a level of novice rather than an expert. However this system provides a flexible computerized work environment for engineers which make acquisition of human expert much easier. Besides, design documents are basis of various works such as maintenance, operation, education etc. and this system is expected to play the role of powerful platform for integrated knowledge base.

b.Diagnosis Support Expert System

In power generation stations major control systems are designed with double or triple redundancies and malfunction of single electrical component does not seriously affect the system. Effects of malfunction may be observed as improper readings of indicators or warnings from monitoring system and failed components need to be replaced. In many cases the effects of failure are deformed through propagation and it is not always easy to pin point a particular electrical element for replacement. Expert engineers inspect design drawings or circuit diagrams and diagnose the system from observed symptoms like human doctors. However compared to human diseases, malfunction of electrical components result in completely different symptom depending on structure of system they belong. As a result in electrical component failure diagnosis, relation between observed symptom and cause are not always as clear as in case of human diseases and engineers rely more on logical reasoning than on experiences or heuristics.

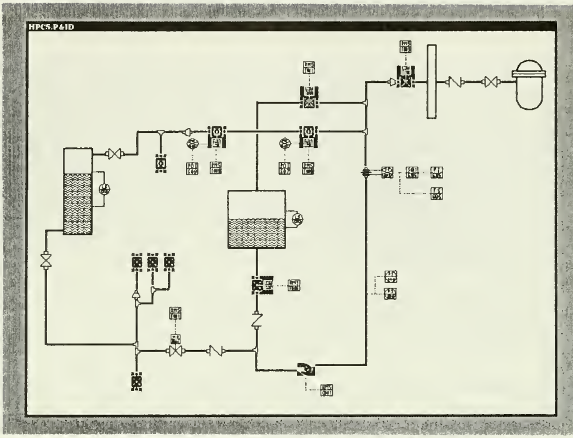


Figure 3. HPCS P&ID Model Display
of Design Support Expert System

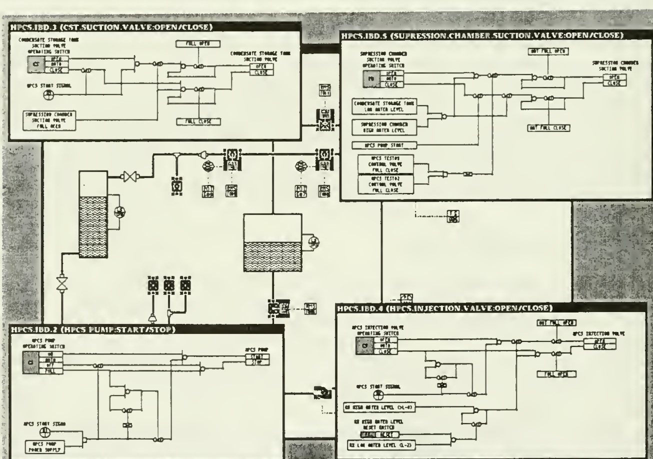


Figure 4. HPCS IBD Model Displays
of Design Support Expert System

Model based reasoning capabilities of PLEXSYS is considered most suitable for these types of problems and an expert system to support diagnosis of electrical component failure in plant control systems was developed. This system use functional block diagrams of control system as domain model (Figure 5) and performs both qualitative and quantitative diagnosis in sequence to decrease the number of suspects and finally points out an element to be replaced. For diagnosis, the system initially uses observed information like indicator readings or monitor outputs. In case the observation is insufficient to figure out single component, the system can optionally make use of additional measurement data like tester readings for further diagnosis. The system was developed on Symbolics workstation and BWR Primary Loop Recirculation (PLR) flow control system was selected as a test case.

EVALUATION RESULTS

As described previously, two application systems were developed to evaluate capabilities of PLEXSYS. The particular systems were designed with intention to cover technical features of PLEXSYS in as wide range as possible. The Design Support Expert System concentrate on integrating wide variety of design drawings using the model representation capability of PLEXSYS whereas the Diagnosis Support Expert System go deep into single type of design drawings. Also the former was developed on general purpose UNIX workstation on the other hand the latter was on specialized LISP workstation, both with same physical memory size. Following are summary of interim evaluation results obtained through development of the application systems.

(a) Model representation capability of PLEXSYS is flexible enough to handle information in various design drawings of plants such as P&ID, IBD, functional block diagram etc.

(b) Interactive graphical interface of PLEXSYS is adequate for building models of around 1000 to 2000 units but for larger models improvements for creating model more efficiently is encouraged.

(c) Reasoning mechanism of PLEXSYS is powerful and flexible as basis for developing various expert systems, yet to customize the function some LISP/KEE skills are necessary.

(d) Performance of application systems depends on computer hardware, model size and complexity of customized functions. For systems around 1000 to 2000 units response speed was acceptable for interactive decision support.

(e) For development of the described application systems, software productivity enhancement is rated around 3 to 10 in magnitude with current PLEXSYS. This means necessary development time of same sort of system are expected to be 3 to 10 times longer without PLEXSYS.

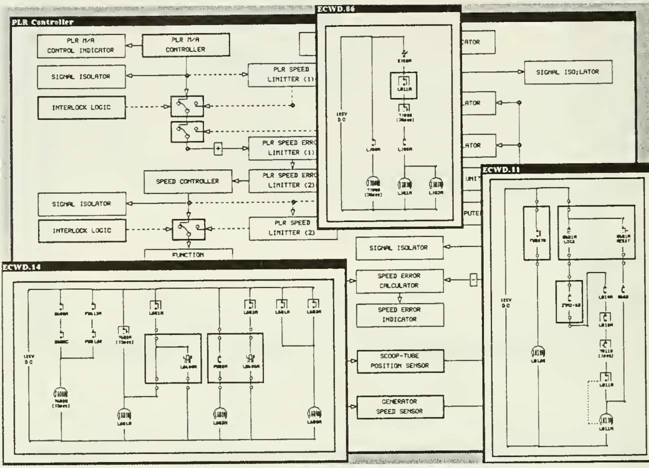


Figure 5. PLR Function Block Diagram Model Display of Diagnosis Support Expert System

(f) In addition to the advantages for individual application system developments, use of common tool allow sharing of domain models and customized functions.

CONCLUSIONS

Under cooperative relation with EPRI, Toshiba participated development of PLEXSYS from early stage. PLEXSYS has gone through its initial stages in laboratory and is on the way towards practical field. Toshiba developed two application systems, design support and diagnosis support expert system to evaluate capabilities and extract necessary improvements of PLEXSYS. Evaluation of PLEXSYS is not yet completed but from the work so far following results were obtained.

The concept of "Model Based Reasoning" can provide powerful solutions to many typical problems in electric power industry and in this point PLEXSYS has great potential to play important role for productivity enhancement and integration of expert systems in this domain. Current capabilities of PLEXSYS is still premature to support engineers willing to use the system without familiarizing themselves to programming. However for engineers interested in developing their own application systems, PLEXSYS already can provide powerful programming environment from both productivity and functionality perspectives.

REFERENCES

1. D.G.Cain, et al., "PLant Expert System (PLEXSYS): Prototype Description", EPRI Report NP-5554, December 1987.

2. H.Sakamoto, et al., "Development of Nuclear Engineering Domain Shell PLEXSYS", Atomic Energy Society of Japan 1989 Annual Meeting Proceedings (in Japanese), April 1989, P.72.

A Knowledge-based System for Heat Exchanger Root-Cause Analysis[1]

REX C. (TRAY) STRATTON and DON B. JARRELL
Pacific Northwest Laboratory[2]
Richland, Washington 99352, USA

Abstract

This paper discusses a software system that provides assistance in the performance of heat exchanger failure root-cause analysis. The system is based on a general model of root-cause analysis. The model was developed from analysis of heat exchanger failures. The software implementation relies on methods and technology developed in qualitative physics and model based reasoning research. Our research leads us to the conclusion that the root-cause analysis process can be modeled, that software systems can and should be developed that implement this process model in an on-line manner, and that root-cause analysis should not, as is current practice, be viewed as a purely reactive analysis but rather as a combination of predictive and reactive analyses.

1.0 INTRODUCTION

This paper discusses a software system that provides assistance in the performance of heat exchanger failure root-cause analysis. The system is based on a general model of the root-cause analysis process. The process model was developed from analysis of heat exchanger failures using structured analysis and artificial intelligence knowledge extraction techniques. The software implementation relies on methods and technology developed in qualitative physics (Bobrow 1985, Hobbs and Moore 1985, Forbus 1988) and model-based reasoning research (De Kleer 1985, Davis and Hamscher 1988). Our research leads us to the conclusion that the root cause analysis process can be modeled, that software systems can and should be developed that implement this process model in an on-line manner, and that root-cause analysis should not be viewed as a reactive analysis but rather as a combination of predictive and reactive analyses.

---

[1] Work supported by the U.S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research.
[2] Operated for the U.S. Department of Energy by Battelle Memorial Institute under Contract DE-AC06-76RLO 1830.

The remainder of this paper is divided into seven major sections: background, approach, process model, qualitative physics, example, and conclusion. Section 2 defines root-cause analysis, discusses why this type of detailed behavior investigation is important, and explains why qualitative physics is used. Section 3 describes our approach for automating this process. Section 4 discusses our model of the root-cause process. Section 5 defines qualitative physics and briefly explains present qualitative physics theories. Section 6 describes the development of the qualitative logic used in heat exchanger analysis. Section 7 provides an example that illustrates our use of qualitative physics. Section 8 summarizes the paper.

2.0 BACKGROUND

2.1 Root-Cause Analysis

We define root-cause analysis as the process of determining the most fundamental cause for process degradation or failure. A cause is labeled as most fundamental if its correction prevents the recurrence of the same process degradation or failure in the same manner. The following example illustrates this definition of root cause.

Suppose while driving a car the driver notices that the engine is overheating and because of this condition decides to stop the car and investigate. An inspection determines that the cause of the overheating is a blown radiator hose. The engine cooling system is subsequently fixed and the blown radiator hose is declared as the root cause. However, after the car is driven another 1000 miles the engine again overheats and the radiator hose is again blown.

This time the driver notifies the car company that he has had the same problem twice. Unknown to the driver the car company has received this same complaint from 50% of the drivers who own cars of this model and year. The car company explains to the driver that the specified radiator hose is not properly designed to operate under the normal cooling system pressure, temperature, and flow. The company has specified a new radiator hose that meets the cooling system design conditions. The new radiator hose is installed in the cooling system and the overheating condition caused by the radiator hose blowout does not recur. The root cause is now properly assigned to the design of the original radiator hose.

2.2 Motivation for Analysis

Nuclear power plants are large complex systems designed to provide safe and cost efficient electricity via the conversion of nuclear energy to electrical energy. These plants require a cadre of highly trained personnel to maintain the plant state consistent with required plant operation and maintenance objectives. Operators continually analyze and determine the state of major components and adjust their behavior to provide the desired overall plant state. Additionally, there are requirements for a technical support organization of engineers and maintenance analysts to identify and characterize expected component degradation. The operators, engineers, and maintenance staff combine their plant knowledge and talents to identify causal mechanisms for degradation and subsequently return these components to their required operability levels.

The function of maintenance is to identify, measure, and correct the degradation and failure phenomena. The performance of maintenance involves a balance between predictive, preventive, and corrective maintenance activities. The balance between corrective or reactive maintenance (repair after failure) and predictive/preventive maintenance (repair before failure) for non-nuclear power plants has traditionally been dictated by operating economics. The cost of component replacement specifies how carefully component performance is monitored and degradation state determined. For nuclear power, safety dominates economics since the potential for a significant impact on the safety of the general public due to component malfunction is dramatically increased. This safety issue coupled with the cost of replacement power for a shutdown nuclear plant (typically $1 million per day) bias the maintenance towards the predictive and preventive maintenance philosophy.

The analysis of degradation mechanics, their impact on component performance, and strategies for correction and mitigation require the coordination of knowledge from all plant operation and maintenance staff. The task of accurate detection, diagnosis, and mitigation requires detailed knowledge of the process physics, materials, and environment. As the plant ages the number of degrading components increases and the ability of the plant staff to determine the complete set of degrading components in a timely manner tends to decrease. This situation results in many ineffective maintenance solutions. It takes the plant staff out of the desired predictive mode and places them in a reactive mode.

We believe that continuous on-line analysis of component degradation could be provided if software systems can be developed that perform the appropriate analysis. These systems must be able to reason about the plant state in the context of goal commands, physical reality, and resulting performance (Seeman, Colley, and Stratton 1983, Stratton and Town 1985). This requirement is similar to that discussed by Davis (1988) concerning observed, predicted, and discrepancy states. If software systems are to provide this functionality, they must be capable of effectively communicating with plant staff, i.e. they must be able to discuss their discoveries and conclusions in qualifiable and quantifiable engineering terms.

2.3 Why Qualitative Physics

Forbus (1988) explains the need for qualitative physics in commonsense reasoning. He discusses the modeling, resolution, and narrowness problems associated with the quantitative approach. We discuss the need for qualitative physics from a different perspective that adds to Forbus's motivation for using qualitative physics in commonsense reasoning. Our perspective is based on an analysis of knowledge requirements for plant operations and an evaluation of how this knowledge is used in problem solving.

If one examines training programs for nuclear operators, it becomes apparent that these programs are founded on physics, mathematics, chemistry, and engineering. The operator is instructed in these disciplines in both a general and plant-specific sense. The operator is then expected to abstract this quantitative knowledge and combine it with the appropriate plant specific knowledge to develop a combination of qualitative and quantitative models necessary for plant operation and maintenance. Armed with these qualitative and quantitative models the operator becomes the principal on-line diagnostician. The extent to which the operator develops and couples these models determines how effective he or she is as an on-line diagnostician.

We view the development of plant/process qualitative models and the integration of these models with quantitative models as necessary for the development of software systems that can predict or diagnose plant degradation at the level needed for safe, reliable, and economic plant operation.

3.0 APPROACH

This section briefly discusses our approach to developing a software system that assists in heat exchanger root-cause analysis. Our approach was biased by the understanding that we needed to determine a model of the root-cause analysis process, specify the process knowledge necessary for root cause reasoning, and develop a representation scheme that implements this model and knowledge.

Figure 1 illustrates the development steps in our approach. The first step consisted of identifing process and component physics (quantitative physics) and representing this physics as quantitative expressions. These expressions were then transformed into qualitative physics expressions using the qualitative calculus discussed by De Kleer and Brown (1984).

The final step was to determine the root-cause analysis logic. This logic was determined using the developed qualitative expressions, predicate logic, and knowledge of failure modes and mechanisms. This step provided qualitative logic expressions that were used directly to analyze and determine the failure root cause.

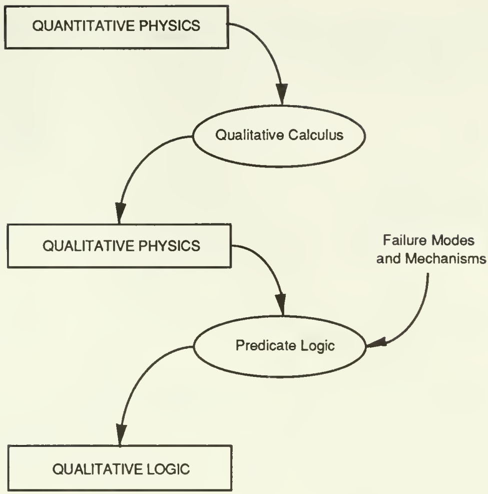


Figure 1 Steps in Qualitative Physics Model Development.

4.0 PROCESS MODEL

A process model must specify reasoning activities, knowledge, structure, and representation. Reasoning activities are transform functions that process information via inference and provide conclusions in the form of facts or requirements. Knowledge consists of the facts, rules, and relations used in the reasoning activities. Structure and representation specify system organization, communication, and control.

Development of the process model was based on the analysis of scenarios of known heat exchanger failures (Jarrell and Stratton 1989). This analysis consisted of selecting and constraining a functionally significant component that has demonstrated recurring failures (Lamb and Leeds 1988). Then a root cause of failure analysis was performed by a system engineer on a number of these failures, which included leaks, blockage, and heat transfer fouling. The systems engineer's analytical process was evaluated. This evaluation resulted in the development of a data-transform model of the root-cause process (Figure 2). To further determine the process knowledge and further develop the representation scheme we augmented the knowledge gained from the analysis of failures with knowledge and concepts learned from a qualitative analysis of heat exchanger physics.

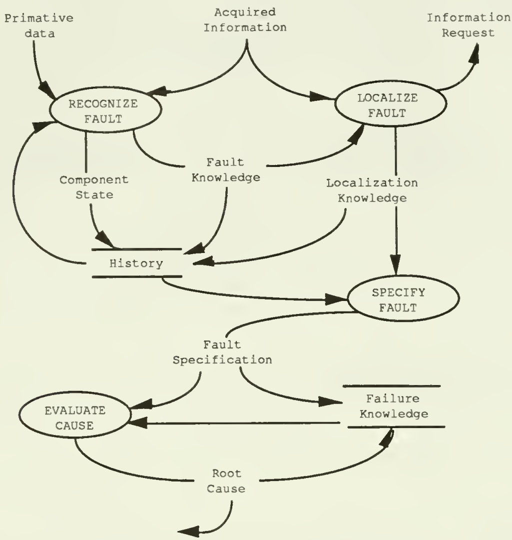


Figure 2 Data-transform Model of the Root Cause Analysis Process.

This paragraph briefly discusses the notation and symbols used in Figure 2. A more detailed discussion can be found in De Marco (1979) or Fairley (1985). Ellipses are used to represent reasoning activities. The activity is described by a strong verb followed by a noun. Thus the fault recognition reasoning activity is described as "recognize fault." Arcs specify information flow (data and knowledge). The direction of flow is indicated by the arrowhead on the arc. Lines without arrowheads indicate that the flow is comming from the reasoning activity through the information descriptor. Therefore, "fault knowledge" is passing from "recognize fault" to "localize fault" and "history." Information descriptors that are inside parallel lines represent information stores.

Our model of the root cause analysis process consists of the reasoning activities: fault recognition, fault localization, fault specification, and root-cause evaluation. The fault recognition activity involves reading (primative data, history, an acquired information), calculating, and comparing information to determine if a fault is going to occur or presently exists. The result of this activity is the development of component and fault knowledge, notification that a fault condition exists, and an activation of further evaluation.

Fault localization processes a wider range of information than fault recognition. The purpose of this activity is to isolate the fault to a specific component and possibly to a subcomponent of the component. This activity may also suggest tasks to be performed for the purpose of acquiring missing information.

The fault specification activity integrates information and conclusions developed in the fault recognition and localization activities to provide a complete description to the fault.

Root cause evaluation is the final activity in the root-cause analysis process. The purpose of this activity is to correlate behavioral discrepancies with potential process disturbances produced by known degradation mechanisms in order that the failure root cause can be determined. The example discussed later illustrates how each of these activities is performed by the system.

5.0 QUALITATIVE PHYSICS DEFINITION

5.1 Definition

A physical system (e.g., the universe, the sun, a chemical processing plant, or a heat exchanger) has a behavior that is determined by its physical properties, structure, and external constraints. Man creates models of physical systems with to better understand their composition and behavior. In order to develop a model, one must first develop a language to represent the model. Integral to the notion of a model is the fact that a model is not the actual physical system but rather an abstraction.

Physical systems can be abstracted in a quantitative or qualitative sense (Kuipers 1986). These abstraction levels are illustrated in Figure 3. Quantitative abstractions model physical systems using the language of quantitative calculus, developed by Newton and Leibnitz, and provide continuous descriptions of the system over time and the real number space. These models become the quantitative physics of the universe, depending on their generality and correctness.

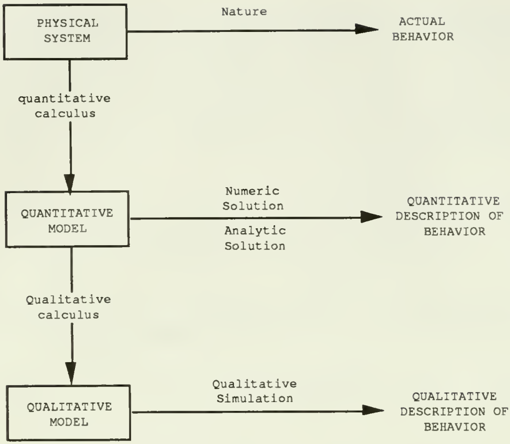


Figure 3 Physical System Abstraction.

Qualitative abstractions model physical systems using the language of qualitative calculus and provide discrete descriptions of the system at discrete instances in time over a qualitative quantity space (Forbus 1988). The quantity space is treated somewhat differently by the various researchers in qualitative physics. We use the quantity space defined by De Kleer and Brown (1984) which reduces the real number space to -, 0, and +.

A formal definition of qualitative physics can be expressed as follows. Qualitative physics is a method of abstraction in which discrete relations that express the qualitative behavior of a continous process are developed.

5.2 An Illustration

The following discusses quantitative and qualitative modeling of fluid mass flow. The quantitative physics describing mass flow of an incompressible fluid in a single phase and constant density is:

$M = \rho v A$	Mass flow rate
$dM/dt = \rho(v\ dA/dt + A\ dv/dt)$	Time derivative

In both equations, each variable has a value in the real number space. These models are used to calculate numeric values of the variables. The equation for mass flow rate is interpreted to mean that the mass flow rate (M) is determined by the product of the fluid density (ρ), the fluid velocity (v), and the flow cross sectional area (A). Solutions to these equations determine the quantitative behavior of the system.

Qualitative physics models relations differently. In qualitative physics we are interested in how the relations relate to the quantity space, i.e. -, 0, +. In general these relations are expressed using operands, operators, and quantity space values (e.g., X(+) or (X-Y)(-)). X(+) means that the value of X is greater than 0 and (X-Y)(-) means that the relation X-Y is less than zero. Solutions to these equations describe the qualitative behavior of the physical system.

For mass flow rate and its qualitative time derivative the qualitative physics expressions are:

M(0)
M(-)
M(+)
(dM - (dA + dv))(0)
(dM - (dA + dv))(-)
(dM - (dA + dv))(+)

5.3 Qualitative Reasoning Theories

A theory for qualitative reasoning must develop qualitative relations, provide qualitative simulation, and be capable of explaining system behavior. Qualitative relations model the physics of the physical system as a function of its structure. Qualitative simulation predicts possible behaviors based on the qualitative relations and initial conditions. Behavior descriptions explain the system behavior based on current values of the qualitative relations.

Presently, there are three different theories used in developing qualitative reasoning systems. De Kleer and Brown (1984) and Williams (1984) develop the relations in terms of components and the paths of interaction provided by connections (device centered ontology). Forbus (1984) develops physical system relations as a function of the processes provided by the physical system (process centered ontology). Kuipers (1986) assumes the qualitative relations are a given and only provides qualitative simulation and behavior description.

Our development of a qualitative model for heat exchanger failure root-cause analysis is based on the device centered ontology.

6.0 HEAT EXCHANGER QUALITATIVE PHYSICS MODEL DEVELOPMENT

Development of heat exchanger qualitative physics is based on the approach discussed in Section 3. Figure 4 is a schematic of the heat exchanger and associated instrumentation. Instrumentation symbols in the figure have the following interpretations: T is temperature and M is mass flow rate.

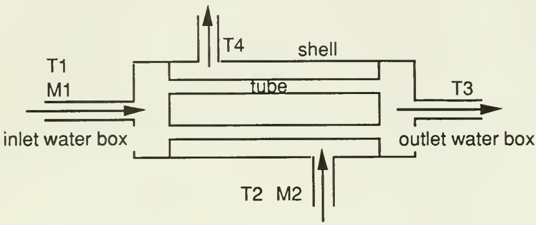


Figure 4 Heat exchanger schematic.

In this section we list the heat exchanger physics and develop the qualitative physics. Additionally, we determine the qualitative logic based on the qualitative physics and knowledge of component failure modes and mechanisms.

6.1 Quantitative Physics

The heat exchanger physics includes conservation of mass flow, conversion of heat energy, mass flow rate, heat changes in a single fluid, and heat exchange between fluids.

$M(in) = M(out)$	conservation of mass
$q(in) = q(out)$	conservation of heat energy
$M = \rho\ v\ Ac$	fluid mass flow
$q = M\ Cp\ \Delta T$	fluid heat change
$qxf = U\ As \cdot LMTD$	heat exchange

In the above equations, q = heat flow, ρ = density, LMTD = log mean temperature difference, v = velocity, Ac = cross section area, As = surface area, Cp = heat capacity of a fluid, and U = heat transfer coefficient across the tubes from one fluid to another.

6.2 Qualitative Physics

To demonstrate how the qualitative physics is developed we will discuss the development of the mass flow qualitative relations. The mass flow equation relates mass flow to fluid density, velocity, and cross sectional area. Of particular interest is the time derivative of this relation, which relates the change in the mass flow to the change in the cross sectional area or velocity.

$dM/dt = \rho(v\ dA/dt + A\ dv/dt)$ quantitative expression

Qualitative relations model the sign behavior of an expression and are not concerned with quantity. Since ρ is constant it is not necessary in the qualitative expression.

$dM/dt = v\ dA/dt + A\ dv/dt$

For variables, the sign signifies the variable's relation to zero, and for derivatives, the sign signifies that the derivative is increasing, decreasing, or constant. Also, dX is shorthand for the qualitative term dX/dt.

$dM = v\ dA + A\ dv$ qualitative expression

The value of the derivative terms in the qualitative expression are either +, 0, or -. Allowing each derivative to take on its allowable values results in the following set of qualitative expressions. Expressions that are not physically realizable (e.g., dM(0) = vdA(+) + Adv(+), are not included).

```
dM(0) = v dA(0) + A dv(0)
dM(0) = v dA(+) + A dv(-), and (v dA) = (A dv)
dM(0) = v dA(-) + A dv(+), and (v dA) = (A dv)
dM(+) = v dA(0) + A dv(+)
dM(+) = v dA(+) + A dv(0)
dM(+) = v dA(+) + A dv(+)
dM(+) = v dA(+) + A dv(-), and (v dA) > (A dv)
dM(+) = v dA(-) + A dv(+), and (v dA) < (A dv)
dM(-) = v dA(0) + A dv(-)
dM(-) = v dA(-) + A dv(0)
dM(-) = v dA(-) + A dv(-)
dM(-) = v dA(+) + A dv(-), and (v dA) < (A dv)
dM(-) = v dA(-) + A dv(+), and (v dA) > (A dv)
```

6.3 Qualitative Logic

The above qualitative expressions and knowledge of failure modes and mechanisms are used to develop logic expressions that imply heat exchanger behavior. Heat exchanger failure (inability to perform designed function) modes consists of leaks (pressure boundary breach), blocks (flow restrictions), and heat transfer coefficient degradation. None of these failure modes affect the velocity directly but rather indirectly through changes in the flow area. Blocks cause the flow area to decrease and leaks act as increases in flow area. The following logic relations model this knowledge (the symbol '=>' is used to signify logical implication):

```
dM(0) => dA(0)    normal behavior
dM(-) => dA(-)    abnormal behavior
dM(+) => dA(+)    abnormal behavior
```

dA(-) => design flow path block or plugging of an existing leak

dA(+) => design flow path leak or dislodging of an existing block

7.0 EXAMPLE

This example illustrates the behavior of the root-cause analysis software and how qualitative relations are used in the analysis of heat exchanger failure conditions. The analysis described by the example is partitioned into fault recognition, fault localization, fault specification, and root-cause evaluation. This example is based on the heat exchanger discussed in Section 6.

The software system is normally interactive. The degree to which the system is interactive is a function of the software system knowledge and the degree to which the component is instrumented for remote data acquisition.

7.1 Fault Recognition

Fault recognition consists of data collection, state calculation, and state evaluation.

	t0	t1
M1	1000	850
M2	833	833
T1	70.0	70.0
T3	90.0	92.0
T2	130.0	130.0
T4	106.0	107.5

Table 1. Heat exchanger sensor data at time t0 and t1.

Data Collection: Sensor data, which describe primitive states, are acquired at specified instances in time, t0,t1,...tn, and stored in a data base. Table 1 gives sensor data at time instances t0 and t1.

State Determination: Higher level component states are calculated using primitive state data and appropriate physics relations. The value of Cp is 1.0 and the sign of the qualitative derivative dM1 is determined by subtracting M1 at t0 from M1 at t1.

```
dM1(-)
q1(t1)  = M1 Cp (T3 - T1) = 1.66 X 10 ** 5 btu/min
q2(t1)  = M2 Cp (T2 - T4) = 1.66 X 10 ** 5 btu/min
qxf(t1) = U As (LMTD) = 1.66 X 10 ** 5 btu/min
```

State Evaluation: The state of each subcomponent is evaluated using facts, relations, and rules. For this example the relavent heat exchanger subcomponents are the inlet water box, outlet water box, and tubes.

The symbol '=>' is used to signify logical implication and the symbol ';' indicates logical or. Facts and implications are recorded as predicate statements. A predicate statement is written as predicate(X,Y); for example, mother(Mary,Ann). A predicate statement is read as 'X _ predicate _ Y'; for example, Mary is_the mother of Ann.

A decreasing value of mass flow rate, dM1(-), is an indicator of abnormal behavior and implies that the flow area has changed in one of the subcomponents (single failure constraint). The software system initates state evaluation in the appropriate sub-components whenever abnormal behavior is determined. A decrease in the cold fluid mass flow rate initiates state evaluation of the inlet water box, outlet water box, and tubes.

Heat exchanger subcomponents affect flow area either through blocking or leaking. If the flow area decreases then either a block has occured in the design flow path or a leak has been patched.

```
due_to(dA(-),wb_in)  => path block; leak block
due_to(dA(-),wb_out) => path block; leak block
due_to(dA(-),tubes)  => path block; leak block
```

At t0 there were no leaks.

```
no_leak(wb_in,t0)
no_leak(wb_out,t0)
no_leak(tubes,t0)
```

The knowledge contained in the no_leak predicates is combined with the knowledge contained in the due_to predicate clauses and concludes with the following statements that specify that the decrease in flow area is due to path blocking:

```
due_to(dA(-),wb_in)  => path block
due_to(dA(-),wb_out) => path block
due_to(dA(-),tubes)  => path block
```

It was determined in the state determination activity that the heat lost by the hot fluid is equal to the heat gained by the cold fluid which is equal to the heat transfered between the fluids.

$$q1(t1) = q2(t1) = qxf(t1)$$

The heat balance fact provides us with no more information about the state of the water boxes. However, the fact that the heat balance is correct does imply that the block is not in the tubes.

The above set of facts resolve to the following statements about subcomponent state:

```
state(wb_in,block); state(wb_in,normal)
state(wb_out,block); state(wb_out,normal)
state(tubes,normal).
```

7.2 Fault Localization

Fault localization analyzes facts in order to localize the cause of the off-normal condition or failure. If there is insufficient knowledge to localize the cause, then recommendations are made which when implemented should provide the missing knowledge. Presently there is not sufficient knowledge to localize the fault. It is known that either the inlet water box is the cause of the fault or the outlet water box is the cause. Because of the ambiguity of fault cause the software system determines that a recommendation must be made. A recommendation is made to inspect the water boxes. The inspection verifies blockage and determines that the blockage is due to clam growth in the inlet water box.

7.3 Fault Specification

The fault can now be specified. The component mass flow decrease at time t1 is caused by blockage in the inlet water box. The blockage is due to clam growth. This new knowledge is logged into the system and associated facts are updated:

```
state(wb_in, block)
block(wb_in, due_to(clams))
state(wb_out, normal).
```

7.4 Root-Cause Evaluation

The root cause of the biofouling can be attributed to design or operation. This is an example of a design root cause because the design environment should be such that in all modes of operation clams cannot grow in the heat exchanger. The root cause can also be attributed to operation if the operation of the heat exchanger specifies that the heat exchanger be thermally backwashed on a periodic basis and that this operation had not been performed as specified.

8.0 SUMMARY

In this paper we discussed a software system that provides assistance in the performance of heat exchanger failure root-cause analysis. The system is based on a general model of the root-cause analysis process. This model was developed from an analysis of the manual performance of root-cause analysis on known heat exchanger failures, knowledge of root-cause mechanisms, and a study of qualitative physics and model based reasoning research. The software for this system is in the process of being coded.

This research leads us to the conclusion that the root cause analysis process can be modeled, that software systems can and should be developed that implement this process model in an on-line manner, and that root cause analysis should not be viewed as a reactive analysis but rather as a combination of predictive and reactive analyses.

9.0 REFERENCES

Bobrow, D.G. 1985. Qualitative Reasoning About Physical Systems. The MIT Press. Cambridge Massachusetts.

Davis, R. and Hamscher, W., 1988. Model-based Reasoning: Troubleshooting. In *Exploring Artificial Intelligence.* Morgan Kaufmann Publishers, Inc. San Mateo, Ca.

De Kleer, J. and Brown, J., 1984. A Qualitative Physics Based on Confluences. In *Artificial Intelligence 24: 7-84.* Amsterdam.

De Kleer, J., 1985. AI Approaches to Troubleshooting. In *Artificial Intelligence in Maintenance.* Edited by Richandson, J.J. Noyes Publications. Park Ridge, New Jersey.

De Marco, T., 1979. Structured Analysis and System Specification. Yourdon Press. New York, New York.

Fairley, R.E., 1985. Software Engineering Concepts. *pp 152.* McGraw-Hill. San Francisco, Ca.

Forbus, K.D., 1988. Qualitative Physics Past, Present, and Future. In *Exploring Artificial Intelligence.* Morgan Kaufmann Publishers, Inc. San Mateo, Ca.

Forbus, K.D., 1984. Qualitative Process Theory. In *Artificial Intelligence 24: 85-168.* Amsterdam.

Hobbs, J.R. and Moore R.C. 1985. Formal Theories of the Commonsense World. Ablex Publishing Corporation. Norwood New Jersey.

Jarrell, D.B., and Stratton, R.C., 1989. Development of a Heat Exchanger Root Cause Analysis Methodology. In proceedings *1989 ANS Topical Meeting: Reactor Operating Experience Plant Operations: The Human Element.* ANS publication.

Kuipers, B., 1986. Qualitative Simulation. In *Artificial Intelligence 29: 289-338.* Amsterdam

Lamb, P.S. and Leeds, E.J., 1988. Operating Experience Feedback Report - Service Water System Failures and Degradation, NUREG-1275 Volume 3.

Seeman, S.E., Colley, R.W., and Stratton, R.C., 1983. Optimization of the Man-Machine Interface for LMFBRs. In *Nuclear Safety Vol. 24, No. 4*. National Safety Information Center at Oak Ridge National Laboratory.

Stratton, R.C. and Town, G.G., 1985. An Artificial Intelligence Program in a Computer Application Supporting Nuclear Reactor Operations. In *Proceedings of the American Nuclear Society International Topical Meeting on Computer Applications for Nuclear Plant Operation and Control.* Published by the American Nuclear Society, La Grange Park, Illinois.

Williams, B.C., 1984. Qualitative Analysis of MOS Circuits, In *Artificial Intelligence 24: 281-346.* Amsterdam.

ELECTRICAL SYSTEMS APPLICATIONS

Knowledge-based System for Voltage and VAR Dispatch

E. DUKE TWEED
Research and Development
Union Electric Company
St. Louis, Missouri 63166, USA

and Electrical Engineering Department
University of Missouri-Rolla
Rolla, Missouri 65401, USA

ABSTRACT

The overall objective of this research effort is to develop a demonstration expert system applied to the control of an electric utility system. This expert system will provide advice in the form of suggested plans of action to be taken to achieve specific goals. The goal is the development of a volt/VAR dispatch expert system which will include the capability of relieving overloaded devices. This expert system utilizes the PROLOG language.

A realistic model of an electric utility system and its interconnections is used in this study. This involves a 630 bus model of the Union Electric Company and its interconnections. This provides an environment in which the results of the expert system can be evaluated and compared with the actions that would be taken in the control center if similar problems occurred. The EPRI power flow program (EPRI EL-599, RP 745) was utilized for the electrical system simulation. Decisions reached in the expert system are passed to the power flow program. The voltage and current profiles are returned to the expert system and the process is repeated until all problems are solved or no further action is possible.

The pattern and amount of generation to be shifted to relieve an overloaded device can be found in a manner consistent with the operation of a control center. The maintenance of a desirable voltage profile is achieved by switching capacitors and reactors and by dispatching VARS from generation buses. The results of this action compare favorably with the action taken in a control center. The major problem with this expert system is the large amount of time required to develop a final plan of action.

Introduction

Expert control using knowledge-based systems is one approach to improving the operation of an electric utility as the systems limits are approached due to the emphasis being placed on greater utilization of the existing generation and transmission system. In addition, the lower amounts of new generation and transmission facilities becoming available in the 1990's will place additional demands for improved control.

The overall objective of this research effort is to develop a demonstration expert system applied to the control of an electric utility system which will be able to provide advice to the operator when disturbances to the system have occurred. This advice will be in the form of suggested plans of action to be taken to achieve specific goals. The goal is the development of a volt/VAR dispatch expert system which will include the capability of relieving overloaded devices. This will be accomplished by switching capacitors and reactors, dispatching VARS from generation plants and by shifting the real and/or reactive generation mix. A realistic model of an electrical utility system and its interconnections is to be utilized in this study so that the results obtained can be evaluated in terms of the actual operation of an electric utility control center.

A knowledge-based system is a computer program that is capable of solving problems hat require expert knowledge in a particular domain. For this study the domain of application is the electrical system and its interconnections. The knowledge base comprises the knowledge that is specific to the electrical system. This includes simple facts about the electrical system, methods, rules of thumb, and ideas for solving problems in this area. Rules of thumb are methods and plans developed through experience. Built into the knowledge-based system is an inference mechanism which provides the means for the system to search for a solution. In this study, the PROLOG language is utilized. PROLOG utilizes a backward reasoning inference mechanism. In backward reasoning the system searches through a collection of facts and rules in order to support a given goal.

There have been two previous knowledge-based systems developed for volt/VAR dispatch. Lui and Tomsovic, "An Expert System Assisting Decision-Making of Reactive Power/Voltage Control" (1), developed this expert system in the OPS-5 language. OPS-5 utilizes a forward reasoning mechanism in which the system looks at a set of facts and rules, and then attempts to reach conclusions about them. This knowledge-based system was designed to correct voltage problems in the electrical network. It was applied to the IEEE 30 bus model.

Tweed developed a demonstration volt/VAR dispatch knowledge-based system in the PROLOG language (2). A realistic model of the Union Electric Company system and its interconnections was utilized. Rules were written to describe the logic sentence that would be utilized to maintain a desirable voltage profile. The PROLOG knowledge-base was linked to a power flow program in order to provide a simulation of the electrical system. Decisions reached in the PROLOG program were passed to the FORTRAN power flow program. The voltage and current profile were passed back to the PROLOG program. This process was repeated until all existing problems have been alleviated. Decisions reached by the expert system were reached in a manner consistent with the operation of a control center.

The Electrical System Simulation

A realistic model of an electrical utility system and its interconnections is utilized in this study. This is necessary so that the results of the knowledge-based system can be compared and evaluated with respect to the results of a control center operator's action if a similar problem occurred in the system under control. The electrical system is modeled with the system in a normal state at peak load. The system is then altered to model realistic problems which could occur. A separate model is developed in order to study problems that could occur under lightly loaded conditions. In an on-line situation, this is unnecessary since the data describing the current state of the electrical system is readily available.

A 630 bus model of Union Electric and its interconnections is utilized in this study. This consists of a 330 bus model of the Union Electric facilities and a 300 bus representation of surrounding systems on 5 of the 7 NERC regional coordinating councils. This model is similar in size to the one utilized for the on-line power flow program at the Union Electric Company control center. For the knowledge-base, all information must be entered in a list format. The generation data included bus name, rated voltage (p.u.), bus type, real generation (MW), reactive generation

(MVAR), maximum reactive generation, minimum reactive generation, maximum real generation, minimum real generation and a weighting factor for economic choice. Bus data, line data and all other needed information describing the electrical system is entered in this manner.

As this knowledge base has evolved, more efficient methods have been found to decrease the time required to update the PROLOG knowledge base. A complete update procedure must be completed on the voltage and current profiles on all buses and lines in the system under control after the power flow program is executed. The backtracking search strategy utilized by the PROLOG language is very inefficient for this process. There is an entry for voltage in each of the 330 bus data descriptions. The volt/VAR program selects one new bus and voltage parameter and searches the bus data knowledge base for a match on bus name. Then this complete entry is deleted and a new one added. This is in sharp contrast to the FORTRAN "DO" loop process of replacing a value in an array. To avoid this problem, the voltage and current profiles are written to disk files in a list format during the report formatting routine of the power flow program in a form compatible with the PROLOG language. When control returns to the PROLOG program, the entire voltage and current profiles are deleted with one command and a load command is executed for the new disk files. Both steps, kill and load, are fast, efficient processes. This also eliminates the need for preparing extensive files before developing the rules to control the system.

Methodology to Remove a Device Overload

The methodology to relieve an overloaded device is listed below. The plan of action is designed to relieve the most severe overloaded condition nearest a generation plant first.

- Examine overloads in the higher voltage system first. If an overload exists, is there an overload between this point and the nearest generation source? Add knowledge of this to the knowledge base.

- Produce a list of generation plants and neighboring areas where increasing generation should be avoided.

- Select the generation plants which are the most sensitive to power flow to the overloaded device to decrease generation.

- Produce a list of generation plants and neighboring areas which are the least sensitive to power flow on the overloaded line for the possibility of increasing generation.

- Determine the amount of generation that needs to be shifted.

- Determine if splitting a bus would be of value in alleviating an overload. If the answer is yes, query the operator to see if this action is to be executed.

- If there is sufficient generation available to accommodate the amount of generation needed, shift the amount of generation obtained in the fifth step from the plant selected in the third step to the plants selected in the fourth step.

- Synthesize all of the plans of action for relieving overloads into a single plan.

- Execute the plan of action for relieving the overloaded devices.

- o Write the results to a data file. Link a FORTRAN program to update the power flow data base. Execute the power flow and pass the results back to the knowledge base.
- o Check the results of the above action. If an overload still exists, repeat the above step.
- o If no overloaded devices are found, link the volt/VAR dispatch section of the program to check the voltage profile.

The process of scanning the overloads nearest the generation buses is designed to relieve as many overloads on the first iteration as possible. The decision to split a bus is based upon an analysis of the line flows in the overloaded substation. Given an unbalance in line flows in the substation, it can be readily determined if the opening of a breaker would be of value.

Results of the Overloaded Devices Program

For this example the system is at peak load. The generation plant on bus 144 has been derated from 285 MW to 165 MW. In addition, the breakers on three 345 kV transmission lines were opened. The net interchange has now changed from 35 MW to -85 MW.

Initially there is some dialogue with the control center operator (Table 1). The response of the operator to the knowledge-based system are underlined.

Table 1

THE INITIAL INTERACTION WITH THE OPERATOR

Is this a continuation of an unfinished job?
No
There are 1740 lines in the normal case.
Enter the number of lines.
1737
The deviation of the net interchange of our area is greater than 100 MWs.
This disturbance is caused by losing generation on Generation Bus 144 by 120 MWs inside our area.
Does this disturbance lead to any losses of a device inside our area?
Yes
Is there any loss of a transmission line inside our area?
Yes
Which transmission line is outaged?

From	To	CKT No.
93	331	1
112	138	1
138	332	1

Is there any loss of a bus inside our area?
No

A plan of action is now developed for balancing the load and generation within the electrical system (Table 2).

Table 2

A PLAN OF ACTION FOR ACCOMMODATING THE LOCAL GENERATION CHANGES

The plan of action for absorbing the deviation of the net interchange is as follows:

Increase generation on Generation Bus 232 by 20 MWs.
Increase generation on Generation Bus 172 by 100 MWs.
Do you want to check the updated data file?
No

After a power flow program has been executed and the results passed back to the knowledge-based system, the electrical system is surveyed for overloaded conditions and high and low voltage problems (Table 3).

Table 3

PROBLEMS REMAINING IN THE SYSTEM

Find out all possible problems within our area.

Overload on Line from 239 to 241 CKT No. 1 by 61 MVAs
Overload on Line from 240 to 335 CKT No. 1 by 65 MVAs
Undervoltage on Bus 144 by 0.0109 p.u.
No bus is overvoltage.

The loss of three transmission lines from a major substation produced an overload on the two transformers at a substation. Two minor voltage problems also existed. The knowledge-based system now searches for the proper pattern to shift generation (Table 4).

Table 4

THE PLAN OF ACTION ON THE FIRST ITERATION

The plan of action for relieving the overloaded line from 240 to 335 CKT No. 1 is as follows:

Decrease generation on Generation Bus 112 by 196 MWs.
Increase generation on Generation Bus 249 by 25 MWs.
Increase generation on Generation Bus 234 by 171 MWs.

The plan of action for relieving the overloaded line from 239 to 241 CKT No. 1 is as follows:

Decrease generation on Generation Bus 112 by 192 MWs.
Increase generation on Generation Bus 234 by 29 MWs.
Buy generation from Area 2 by 163 MWs.

The plan of action for relieving all of the overloaded lines is as follows:

Increase generation on Generation Bus 249 by 25 MWs.
Decrease generation on Generation Bus 112 by 388 MWs.
Increase generation on Generation Bus 234 by 200 MWs.
Buy generation from Area 2 by 163 MWs.
Adjust the scheduled net interchange to -128 MWs.

On the second iteration it was found that the overload on the transformers had been reduced by one-half (Table 5).

Table 5

PROBLEMS EXISTING ON THE SECOND ITERATION

Find out all possible problems within our area.

Overload on Line from 239 to 241 CKT No. 1 by 32 MVAs.
Overload on Line from 240 to 335 CKT No. 1 by 33 MVAs.

Undervoltage on Bus 144 by 0.01 p.u.
Overvoltage on Bus 156 by 0.0068 p.u.
Overvoltage on Bus 234 by 0.0098 p.u.

A second plan of action is now developed to deal with the remaining overloaded conditions (Table 6). This program can be stopped at this point and restarted at a later time if desired.

Table 6

PLAN OF ACTION ON THE SECOND ITERATION

The plan of action for relieving the overloaded line from 240 to 335 CKT No. 1 is as follows:

Decrease generation on Generation Bus 112 by 66 MWs.
Decrease generation on Generation Bus 215 by 97 MWs.
Buy generation from Area 3 by 163 MWs.

The plan of action for relieving the overloaded line from 239 to 241 CKT No. 1 is as follows:

Decrease generation on Generation Bus 112 by 0 MWs.
Decrease generation on Generation Bus 215 by 162 MWs.
Buy generation from Area 4 by 162 MWs.

The plan of action for relieving all of the overloaded lines is as follows:

Decrease generation on Generation Bus 112 by 66 MWs.
Decrease generation on Generation Bus 215 by 259 MWs.
Buy generation from Area 4 by 162 MWs.
Adjust the scheduled net interchange to -453 MWs.

Again line flows are compared with emergency ratings and the problems are listed in Table 7.

Table 7

REMAINING PROBLEMS

Find all possible problems within our area.

No line is overloaded.
Undervoltage on Bus 144 by 0.0092 p.u.
Overvoltage on Bus 156 by 0.0076 p.u.
Overvoltage on Bus 234 by 0.011 p.u.

Since there are no remaining overloaded devices, the previous plans of action are merged into one final plan of action (Table 8).

Table 8

FINAL PLAN OF ACTION

The final conclusion for the plan of action to deal with this contingency is as follows:

Increase generation on Generation Bus 232 by 20 MWs.
Increase generation on Generation Bus 172 by 100 MWs.
Increase generation on Generation Bus 249 by 25 MWs.
Increase generation on Generation Bus 234 by 200 MWs.
Decrease generation on Generation Bus 112 by 454 MWs.
Decrease generation on Generation Bus 215 by 259 MWs.
Buy generation from Area 2 by 163 MWs.
Buy generation from Area 3 by 163 MWs.
Buy generation from Area 4 by 162 MWs.
Adjust the scheduled net interchange to -453 MWs.

The process of initializing the knowledge base requires the execution of two power flow programs. The first time the electrical system is modeled in a normal state and the second time the electrical system is altered in order to represent problems requiring attention. In an on-line situation in a control center, the above two power flow program executions would not be necessary. Actual data would be available from the System Control and Data Acquisition System or the state estimator. The decisions reached to solve the problems in this section are realistic and consistent with the operation of a control center.

Volt/VAR Dispatch

The volt/VAR dispatch section of this expert system is designed to maintain a predetermined voltage profile in the electrical system. This objective is met by switching controllable capacitors and reactors and by raising or lowering the voltage at a generation bus. Under certain conditions a transmission line will be taken out of service to relieve high voltage problems. The principal actions to be taken are listed below.

- o Examine the voltage profile at all generation buses. Adjust the voltage by raising/lowering voltage.

- o Examine all points of interconnection. Switch capacitors or dispatch VARS from generating plants.

- o If the previous step fails, request assistance from adjoining utility.

- o Examine all load buses. If the voltage is high, determine if the state was reached by a previous action.

- o If the answer is "yes" to the previous step and the problem is serious, cancel previous action and find a new alternative.

- o If no other alternative exists, inform the system operator

- o Switch capacitors off and/or decrease VAR flow from the appropriate generating plant.

- o If the voltage is low at a load bus, repeat the equivalent actions to be taken in the previous three steps.

- o If the system load is low and the voltage profile in the 345 KV transmission system is above normal, consider taking a long transmission line out of service, if that line is lightly loaded.

- o If no other alternatives exist, inform the system operator.

In this example, problems are created so that high and low voltage problems existed throughout the electrical system. Capacitor banks which should have been switched on are switched off. VAR flow from generation plants is not sufficient to bring the voltage up to an acceptable level at some load buses. A power flow program is executed with the data base altered to represent the sample problem. The knowledge base is then updated with the results of this action. The following voltage problems are then identified (Table 9).

Table 9

INITIAL PROBLEMS FOR VOLT/VAR DISPATCH

The voltage on Bus 30 is low 0.8729 p.u.
The voltage on Bus 39 is low 0.9415 p.u.
The voltage on Bus 68 is low 0.9651 p.u.
The voltage on Bus 98 is low 0.9635 p.u.
The voltage on Bus 111 is low 0.9592 p.u.
The voltage on Bus 123 is low 0.9253 p.u.
The voltage on Bus 220 is low 0.9311 p.u.
The voltage on Bus 251 is low 0.9356 p.u.
The voltage on Bus 290 is high 1.0494 p.u.
The voltage on Bus 302 is low 0.9441 p.u.
The voltage on Bus 308 is low 0.9636 p.u.
The voltage on Bus 310 is low 0.9502 p.u.
The voltage on Bus 323 is low 0.9258 p.u.
The voltage on Bus 324 is low 0.9418 p.u.
The voltage on Bus 325 is low 0.9457 p.u.

There is a capacitor bank at Bus 290 which can be taken out of service. There are capacitor banks at Buses 39 and 251 which can be switched on for VAR support. The first action will be to switch all capacitor banks which have the potential of improving the voltage profile (Table 10).

Table 10

SWITCH CAPACITOR BANKS

Increase caps. on Bus 39 by 33.6 MVAR
Increase caps. on Bus 251 by 24.3 MVAR
Increase caps. on Bus 290 by -6 MVAR

There is a capacitor bank at Bus 290 which can be taken out of service. Therefore capacitor banks at Buses 39 and 251 can be switched on for VAR support. The first action will be to switch all capacitor banks which have the potential of improving the voltage profile (Table 11).

Table 11

VOLTAGE PROBLEMS AFTER SWITCHING CAPACITORS

The voltage on Bus 30 is low 0.8727 p.u.
The voltage on Bus 68 is low 0.9708 p.u.
The voltage on Bus 123 is low 0.9645 p.u.
The voltage on Bus 220 is low 0.9731 p.u.
The voltage on Bus 290 is high 1.0474 p.u.
The voltage on Bus 302 is low 0.9461 p.u.
The voltage on Bus 308 is low 0.9664 p.u.
The voltage on Bus 310 is low 0.9802 p.u.
The voltage on Bus 323 is low 0.9672 p.u.

Voltage problems on Buses 39, 111, 241, 251, 324, and 325 have been eliminated by switching capacitors. The next action is to dispatch VARS from the generating plants (Table 12).

Table 12

DISPATCH VARS FROM GENERATING PLANTS

Increase voltage on Bus 28 by 0.01 pu
Increase voltage on Bus 172 by 0.01 pu

The results of this action show that the voltage problem at Bus 30 is eliminated (Table 13).

Table 13

PROBLEMS REMAINING AFTER DISPATCHING VARS FROM GENERATING PLANTS

The voltage on Bus 68 is low 0.9740 p.u.
The voltage on Bus 123 is low 0.9657 p.u.
The voltage on Bus 220 is low 0.9744 p.u.
The voltage on Bus 302 is low 0.9475 p.u.
The voltage on Bus 310 is low 0.9533 p.u.
The voltage on Bus 323 is low 0.9685 p.u.

An adjoining utility is in a position to dispatch VARS for support of buses 123, 302, 308 and 310. The change in the voltage level on Buses 68 and 220 by the above action is not sufficient to warrant further VAR dispatch. After establishing a problem for the volt/VAR dispatch exert system to solve, the load flow program is executed two times. In this example, it would be desirable to switch capacitors and dispatch VAR flow from generation plants on the same iteration. It is a relatively straight forward process to dispatch VARS from the generation plants. The results of this simulation are reasonable and consistent with the operation of a control center.

Conclusions

Knowledge-based systems organized to serve as a control center operator's assistant have been shown to have the potential of being a feasible and valuable asset to improving the operation of an electric utility. The pattern and amount of generation to be shifted to relieve an overloaded device can be found in a manner consistent with the operation of an electric utility control center. In like manner, it was shown that the pattern and amount of reactive power required to provide a desirable voltage profile can be found with this approach. However, there are problems that have to be solved before this can become a reality.

This knowledge-based system approach does not rely on the prior development of contingency plans. Typically in a control center, contingency plans are available to an operator which have been developed with the use of power transfer distribution factors. Most single contingency problems, which are of significance, are analyzed in this manner. It is not possible to analyze all multiple contingency problems which could occur. One of the important attributes of the knowledge-based system approach is that the number or pattern of outages occurring is not significant. This knowledge based system will only fail when the power flow program fails to find a solution.

The value of knowledge-based systems applied to electric utility system control will increase as the system operation grows in complexity. This situation could occur as greater emphasis is placed on utilizing existing facilities and also due to the lack of new generation becoming available in the 1990's. Control center operators need little assistance with single contingency problems. Multiple contingency problems demonstrate the need for an operator's assistant. An overloaded device situation that is confined to a limited area does not present a difficult problem to the control center operators. An example of a situation in which a knowledge base can be of value is where overloaded devices exist at several points throughout the electrical system. The process of shifting generation from one plant to another may alleviate the problem in one area and aggravate it at another.

The major problem with this knowledge-based system is the large amount of time required to provide advice. The time required to provide the operator with advice is limited to a very few minutes. This knowledge-based expert system cannot respond in that time frame. To be used in a realistic manner, the power flow program should be linked one time. This means that the amount of generation to be shifted to relieve overloaded devices and voltage problems must be calculated. The approach used in this study was basically to simulate the effect of ramping generators by changing the real and reactive generation in increments.

Acknowledgements

The research project is sponsored by the Electric Systems Division of the Electric Power Research Institute under EPRI RP 2944-2. Dr. David Curtice is the EPRI project manager.

References
1. C. C. Liu and K. Tomsovic. "An Expert System Assisting Decision Making of Reactive Power/Voltage Control". IEEE Trans. on Power Systems. Vol. PWRS-1, No. 3, Aug. 1986, pp. 195-201.

2. E. Duke Tweed. "Knowledge-Based Systems for Voltage and VAR Dispatch". Final Report. EPRI RP 2944-2. May, 1989.

Load Control Expert System

PAUL EDMUNDS
Duke Power Company
Information Systems Department
P.O. Box 33189
Charlotte, North Carolina 28242, USA

ABSTRACT

Duke Power's load control system is designed to interrupt electrical power supplied to approximately 200,000 residential water heaters and air conditioners, allowing Distribution Department personnel to shed approximately 400 Mw of electrical load. Two minicomputers in the Charlotte general office communicate through modem connections with approximately 340 Substation Control Units (SCUs) in distribution substations. These SCUs use power line carrier technology to broadcast signals to the residential devices participating in the load control program. Information on the status of the SCUs is gathered on a continuous basis, stored on the Charlotte minicomputers, and used to diagnose communications errors. An expert system was developed to read the status files and report several communication error types. It was developed with Nexpert Object and delivered with the Nexpert Object Run Time (NORT) environment for execution on an IBM PS/2 workstation.

LOAD CONTROL SYSTEM HARDWARE

The load control system consists of a Data General MV 8000 and a Digital Equipment Corporation VAX 11/750 minicomputer located in the Charlotte general office. Each minicomputer communicates via modem and dedicated communication lines with Substation Control Units (SCUs) in approximately 170 distribution substations throughout the Duke Power service area. The SCUs receive control signals for the residential water heaters and air conditioners which they broadcast to these devices using power line carrier technology. The Data General system was chosen for this expert system project because it can report more diagnostic information through a transponder located on one of the busses coming from each substation. This transponder monitors and responds to signals sent from the SCU; these responses are reported back to the central system. This system is diagrammed in Figure 1.

LOAD CONTROL SYSTEM SOFTWARE

Two types of error checks are performed on the communication components of the load control system. In the first error check, a query is sent to each SCU to determine if it is operable. The second error check involves the two-way communication portion of the SCU and a status register in the transponder.

SCU operability (first error type) is determined by an interrogation of each SCU every 15 minutes. If the SCU does not respond to the interrogation, the time and date of the attempted interrogation and the id number of the SCU are written to an error file.

The transponder status (second error type) is determined as follows. A program running on the host computer sends a command in the middle of every hour to each transponder. This command sets the transponder status register to either an "S" or an "H", depending on the hour. At the beginning of each hour, each status register is interrogated and the value found is recorded. If there is a problem communicating with the transponder, then the host determines the error type and this value is recorded instead of the "H" or "S" expected for that hour. Status data are accumulated for a 24 hour period. Therefore a normal file with no errors should read "SHSHSH...SH" for each SCU. Deviations from this pattern are interpreted as communication errors. The following errors can be determined from the patterns:

- o bad communication error
  This error is noted if a "C" is found in the status code string.
- o scram error
  A scram error is indicated if more than 12 "L"s are found in the status code string.
- o device lock
  A string of five consecutive "B"s (i.e. "BBBBB") indicates a device lock.

The status code for each string is scanned for these patterns beginning with the last reading for the day. A device lock error can be noted with any other error, but only one bad communication or scram error can be asserted for any one SCU. An SCU that reports a bad communication error from the status report and is also on the error report for the same time has a two way communication error.

EXPERT SYSTEM APPROACH

An expert system was developed with the following goals:

- o automate scanning of the status reports
- o determine communication errors
- o report the communication errors
- o learn about the technology and development of expert systems

The load control expert system was developed on an IBM PS/2 Model 80 with Nexpert Object software. Several factors influenced the selection of Nexpert. A major requirement was for software that could run on the PS/2 platform without significant hardware enhancements. A system was also wanted that would offer sifnificant function; this was desired both to solve the load control diagnostic problem and to serve as a system to help us learn about the field of expert systems. Nexpert also offered an environment that could be linked to external files.

The load control diagnostic system combines both conventional C-language programs and Nexpert (Figure 2). A C program was developed to "preprocess" the status file; traditional loop logic was determined to be the most efficient way to read through the 24 hours of status values and determine the appropriate error condition. The status code strings for each SCU are evaluated as described in the section on the transponder status checks, and a status output file is created that contains the id for each SCU, the presence or absence of the three error types that can be determined from the status report and the time that the error type (if found) occurred.

Nexpert is then loaded and each SCU becomes an object in the class of SCUs, using the Retrieve and CreateObject actions of Nexpert. Pattern matching rules then pick out the SCUs for each error type, placing them in new classes that are written to external files for reporting. Those SCUs that are found on the error log are read into Nexpert and assigned to a new class. A rule next selects the objects (SCUs) common to this class and the class of bad communication SCUs. The common objects that have errors at the same time are written to a new class representing the objects with two way communication errors. An external file containing these objects and their attributes is created and this file is printed out. The entire expert system consists of only 13 rules. This rule count is low because of the C preprocessor program and the use of object representation and pattern matching.

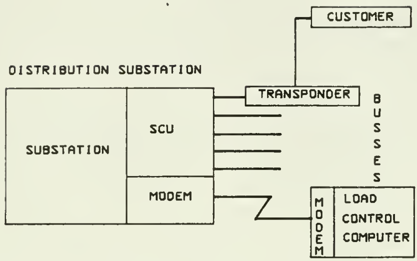
DELIVERY

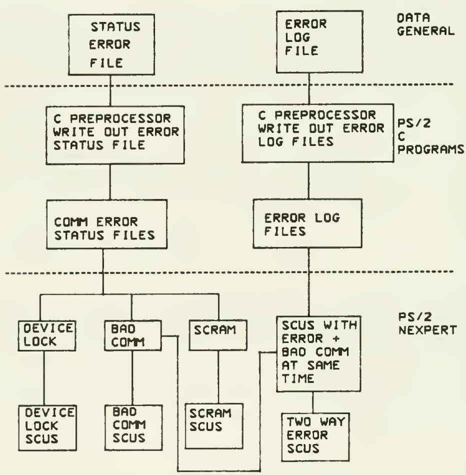
The system was initially prototyped for delivery with a graphics based interactive user interface. However, upon review of the prototype the users stated their desire for a completely "hands off" system requiring minimal user interaction and a printed report. The Nexpert Object Run Time (NORT) environment was investigated and found to meet these requirements, allowing the system to be placed in a DOS BAT file. The user types in the name of the BAT file which executes the C programs and creates the error files. Then the Nexpert Run Time Definition (RTD) file is loaded. The RTD file loads the knowledge base and begins processing the rules, assigning the SCUs to the appropriate error classes and creating the report files. At the conclusion of the knowledge processing, control returns to the BAT file and the report files are printed. After the user types in the BAT file name, no prompting or system monitoring is required.

OPERATIONAL EXPERIENCES

The system has been delivered to the end users and is currently undergoing testing and evaluation. Initial response to the system has been favorable; it clearly meets the requirements for an automated solution for limited error diagnostics.

LOAD CONTROL SYSTEM





One major drawback to the system is its execution speed. An analysis of all SCUs and error conditions takes over 3 hours to complete. If the user did not require an unattended system this would be a fatal problem; in the batch environment it is not as critical to produce results quickly. The long execution time is directly related to the large memory requirements during object creation and the constraints of the DOS environment. Over 2000 objects are created from the error log, as each SCU at a particular time becomes a unique object. NORT is not able to use expanded memory for these objects, so the input file must be split into 4 files. Each piece is processed separately, and the memory is freed before the next file is read in. Reading and writing these files also increases the rule count. When the system is run under the Nexpert Object Development system it runs considerably faster (in about 1 hour) due to the cache software in Microsoft Windows. NORT cannot take advantage of this software.

SYSTEM ENHANCEMENTS

Enhancements to the system fall into three areas: increased error detection, improvements in the execution environment, and better reports. The errors that the system currently detects are a basic set; the load control system is susceptible to more error types. Rules will be added to determine when these occur. This will enhance the system and also test the ability of the system to be modified. The execution environment will be enhanced by decreasing the execution time and automating system execution. A DOS protected mode run time version of NORT should allow utilization of higher memory and may speed up execution. Scheduling the system to run at night will make the execution speed less of a factor if the reports are available at the start of each day. This will also result in a completely automated system. The reports are now generated by simply printing out the Nexpert files written by the system. Processing these files with a report generator will help in the readability of the reports. A C program will be developed to perform this function.

CONCLUSIONS

The goals of this project were to develop an automated system that could scan communication error reports, determine the communication errors, and report these errors while learning about the technology and development of expert systems. These goals have been met in the development of the load control expert system. A usable system has been delivered to the Distribution Department that will help free the human experts from the routine of interpreting error reports. In addition, much has been learned about the development and application of expert systems technology, and this information is being disseminated into the Information Systems Department. The load control system will continue to be refined, and expert system technology will be applied to other areas of the Company.

A Rule-based Load Shedding Strategy in Electric Power System

S. S. SHAH
Electrical Analytical Division
Sargent and Lundy Engineers
Chicago, Illinois 60603, USA

S. M. SHAHIDEHPOUR
ECE Department
Illinois Institute of Technology
Chicago, Illinois 60616, USA

ABSTRACT

This application provides a practical methodology and notion for developing systems capable of knowledge–intensive performance. The AI technology would allow us to develop a procedure in such a way that the task of decision making for a stable operation of a large power system would be performed based on rules and axioms as well as the data pertaining to a particular state of the system. The objective of this study is to develop an expert system which would analyze the security of a large power system in the real time, and help an operator in his critical decision making for the system recovery. The advantage of using this approach versus conventional algorithmic approaches is the fact that an algorithmic approach has to examine the data exhaustively for making any type of computations, whereas expert systems consider rules and select the data relevant to a particular situation and problem. This would limit the computations to mostly affected parameters, and improve the efficiency of the decision making process. Furthermore, the time of the execution does not change significantly with the size of the system, primarily because the corrective action is offered on a local basis. The application of this approach to a 30-bus system is discussed in the paper.

INTRODUCTION

In recent years, advanced automation in power systems has permitted the implementation of more sophisticated energy management systems which allow enormous volumes of data to be handled more rapidly, more reliably and more accurately. These innovations have provided enhanced mechanisms to assess the state of a secured power system. However, one of the main problems associated with the operation of an electric power system is the decision making within a short period of time according to a set of information produced by the power grid upon the detection of a fault. As the size of the system increases, it becomes more and more complicated for an operator to recognize the detailed state of an emergency that would exist in a system and prescribe appropriate responses to restore the normal operation of the system. Any recommendations which could speed up the decision making process and enhance the likelihood that an operator would take only those steps which are in the best possible interest of the continuous, safe, and proper operation of the power system must be seriously taken into consideration. Most of the modern power systems are designed such that they can tolerate almost all major disruptions, however, depending on prevailing circumstances, a dynamic system may not be able to perform satisfactorily and meet system criteria at all times. This is due to the fact that many components may have been taken out of service for maintenance or, have been on forced outages and a power system may not be operated with all the resources in service. Hence, the job of an operator is to try, within economic and design limitations, to maximize the system reliability.

The advantages of implementing an expert system in a complicated decision making process are as follows :

- An expert system would always be available in a control center for a specific application and never retires. So, continuous improvements in its performance is possible.
- Expert system capability will not deteriorate over a long period of time despite the fact that it may perform similar tasks over and over.
- In critical moments of decision making, an expert system will not be affected by the severity of a contingency, environmental conditions, or the number of staff available in the operating room .
- Many expert systems performing different tasks can be integrated into a global system.

The objective in power system security analysis is to keep the system in operation once a contingency has occurred and before its effect has been corrected. Hence it is necessary to consider the effect of adjusting various control components, such as governors and excitation controls, or options such as load shedding as key alternatives in the operation of a power system. Currently, security analysis in energy control centers is tackled by human operators. Decisions made by an operator are based on his experience regarding the operation of a large network, the knowledge that he has acquired based on his conversations with his superiors and power system engineers, his memory to recollect the related information, and the overall set of data which represents various measurements such as voltages, currents, power factors, power flows, etc. Actions that an operator would take in a critical situation, depends largely on the state of his mind. However, it is generally believed that in critical conditions a human being is likely to panic and make irrational decisions, which would cause a greater emergency and eventually a catastrophe.

Major characteristics of a rule-based system that are implemented in the design of a power system security analyzer should fulfill the following criteria:

- Applications of Artificial Intelligence techniques to the control and operation of a large power system, and the identification of a systematic procedure for decision making that an operator would follow in critical circumstances regardless of the type, size and location of faults in a power system.

- Localization of control actions in an emergency situation using a logical reasoning, which will speed up the decision making process and will reduce the required memory space for a very large scale power network. This is quite contrary to numerical algorithmic procedures which have been implemented in the past.

- Selection of the most effective control devices for power system restoration once an emergency has occurred in a network.

- Prioritization of the available control tools in a network for reducing the cost of operation, and the degradation of the system.

SOLUTION TECHNIQUES

The power system security analyzer would facilitate a rational and quick decision making process in a troubled power system. The main objective of this analyzer is to make comprehensive use of sensitivity analyses, distribution factors, and load decrement superposition principle to alleviate overloads in various transmission lines as well as the violation of voltage profile in a power network. Figure 1 represents the scope of the power system security analyzer. The power system analyzer makes use of numerous data such as the real power flow in a transmission line, the voltage magnitude at a bus, etc., as well as the data regarding the topology of the system which is readily provided by data acquisition systems and recorded in energy management centers.

The implementation of expert systems in power systems operation and control covers a wide range of applications. In order to design the analyzer, the following types of contingencies are considered in this study :

- Component overloads

- Bus overvoltages

In critical circumstances, if some of system components are overloaded, various types of control actions would be available to a system operator which could be utilized to reduce line

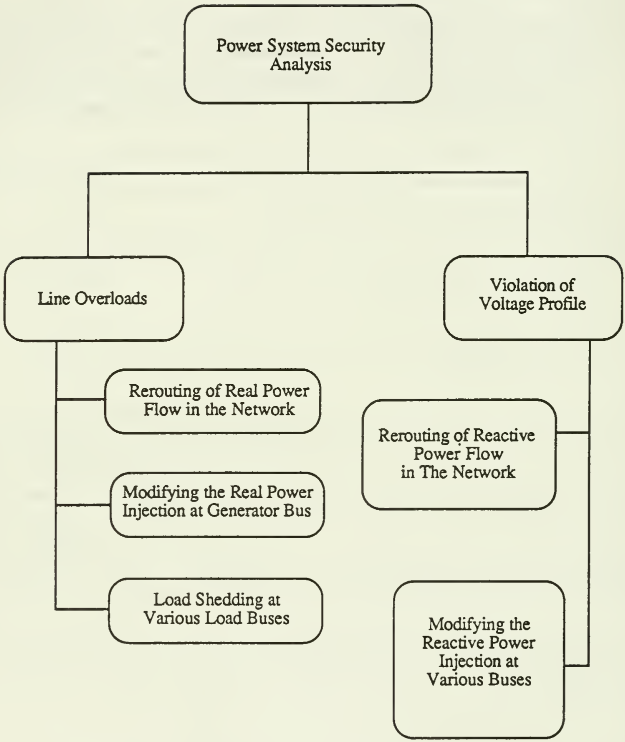


Figure 1 Scope of Power System Security Analyzer

overflows in the network. The following alternatives for reducing component overloads would be considered in this approach:

- Power system emergency control
- Load Shedding

Power system emergency control represents specific remedial actions which would be executed if a contingency occurs in the system. In this regard, following actions would be considered:

- Adjusting the control transformers
- Shifting the real power generation

These remedial actions represent procedures for rerouting real power flows in a system in the given order. So, let's assume some of the existing components are overloaded due to a contingency in the network. To save healthy components in the system, one has to release overloads by transferring flows to transmission lines which are not loaded up to their maximum capacity. In implementing these ideas, following sensitivity factors are provided as inputs to the expert system:

- The change in real power flow in a transmission line due to the change in the real power injection at a generator bus. This sensitivity is termed as a **A** sensitivity.
- The change in real power flow in a transmission line due to the change in the tap-setting of a phase-shifting transformer. This sensitivity is termed as a **U** sensitivity.
- The change in the voltage magnitude at a bus due to the change in reactive power injection at a bus. This sensitivity is termed as a **D** sensitivity.
- The change in the voltage magnitude at a bus due to the change in the tap-setting of a control transformer. This is termed as a **T** sensitivity.

The mathematical derivation of these sensitivity factors is described in references .

Using these values, the most appropriate component in the power network that would require a minimum adjustment for alleviating specific component overloads would be identified. The selection criterion is based on the fact that remedial actions should not cause any additional component overloads in the system. Furthermore, the expert system should try to adjust tap settings of available control transformers for rerouting additional power flows, and if not enough transformers are available in the system, or if available transformers are not located in proper positions, then the expert system would consider the reallocation of real power generations at specific buses in order to reduce the tension in the system. Again, the selection criterion for the most appropriate generating unit will be based on the sensitivity of overloads to various adjustments of the injected real power to the system.

The effectiveness of various procedures for rerouting real power flows depends on the location of overloaded lines, as well as the operating state of a power system. This is due to the fact that very large changes in the power injection may result in very small changes of the real power flow in a remotly located transmission line. Hence as much as possible, adjustments should be done locally. However, due to the existence of various power system constraints and system operating conditions, it is not always possible to adjust the injection locally. For example, a generating unit may not be available at nearby buses, generators at nearby buses may be running at their full capacities, or changes in MW injections at nearby buses may overload other transmission lines in the system.

These factors constitute the selection criteria for rescheduling the MW generation and alleviating overloads in transmission lines. Based on the criteria introduced in this study, the most appropriate generator is selected and its MW generation is altered accordingly. It is always required to review the procedure in order to make sure that in the process of alleviating an overload, other healthy transmission lines in the system would not be overloaded.

If the emergency control fails to restore the normal operation of the system, the expert system considers the load shedding as another alternative for reducing overflows. Figure 2 represents various factors affecting the load shedding scheme. However, in an emergency, the problem associated with an appropriate load shedding schedule for a given contingency and at a given system state must be resolved with extra caution, because an unnecessary load shedding creates unsatisfied customers as well as the loss of revenue to utilities. In order to minimize the required amount of load shed and release overloads in a short period of time, the load shedding scheme will be implemented in two stages which are described as follows,

- First we will make a quick and conservative estimate of the required amount of load shed for the removal of overloads from the system.

- Then, based on the available optimization alternatives and the status of the power system, we will optimize network flows and restore fractions of the load accordingly to satisfy the demand as closely as possible.

The problem of load shedding can also be viewed as the optimum load dispatching problem under abnormal operating conditions. In other words, it represents an optimal load dispatch with additional constraints, which takes into account system abnormalities. To reduce the risk of deterioration of a system due to load shedding, following conditions must be considered: Loads must be dropped temporarily and instantaneously in those parts of the system where the power has become deficient. The load curtailment should be avoided in those parts of the system where a temporary excess of power would cause generators to speed up, and consequently drop out of service. At all times, the generation must be scheduled such that additional power can be produced rapidly and transported to those parts of the system where power has become deficient.

These operations are currently performed by a human operator, based on his experience and his knowledge of the dynamic behavior of the system. responses to restore the normal operation of the system. we would consider a procedure that would accomplish these goals using heuristics. In this regard, a quick estimate of the amount of load that must be shed is determined according to the following two procedures,

- Flow Distribution,
- Load Decrement.

This two procedures are described as follows:

Flow Distribution. Using this procedure we would determine the flow reduction prescribed for each line. Suppose that there are n lines connected to a bus, m out of n lines have power flowing into the bus, the overload in line i is denoted by IL_i, and AF_i is the actual real power flow in line i. From the existing state of the power system, if we would like to decrease the flow in line i by IL_i, we would have to reduce the real power flow in all m lines connected to that bus.

So, the amount of flow that should be reduced in line k is determined by the following equation,

$$IL_k = AF_k \times \frac{IL_i}{AF_i}, \qquad k = 1, \ldots, m$$

where, IL_k is the amount of flow reduction in line k, and IL_i/AF_i is defined as the *overload factor*. If more than one line is overloaded at a given bus, then one has to take the maximum of the respective overload factors as a common overload factor for all the incoming lines. To account for approximations, all the line flow limits are set slightly below its nominal ratings, i.e. 95 % of the actual flow limit.

Load Decrement. Suppose that there are n lines connected to a bus, m lines have power flowing into the bus, $n - m$ lines have power flowing out of the bus, OL_i is the amount of real power flow that is be reduced in lines carrying power out of the bus, and IL_i is the amount of real power that needs to be reduced in the lines which carry power into the bus. Then the incoming overload for the given bus is defined as,

$$incoming\ overload\ = \sum_{i=1}^{m} IL_i$$

and, the outgoing overload is defined as,

$$outgoing\ overload = \sum_{i=1}^{n-m} OL_i$$

$$\begin{aligned} load\ shed &= incoming\ overload - outgoing\ overload \\ &= \sum_{i=1}^{m} IL_i - \sum_{i=1}^{n-m} OL_i \end{aligned}$$

where, the *incoming overload* > *outgoing overload.*

If *outgoing overload* > *incoming overload*, and there is a generator connected to that bus, then the reduction in generation is given by,

$$\begin{aligned} generation\ decrement &= outgoing\ overload \\ &\quad - incoming\ overload \\ &= \sum_{i=1}^{n-m} OL_i - \sum_{i=1}^{m} IL_i \end{aligned}$$

Optimization of Network Flows. Suppose that LS is the amount of load shed at a given bus. There are n lines connected to that bus, out of which m lines have inflow of the power, and l out of $n-m$ lines have reached their power flow limits. So if we can feed the power to this bus through other non overloaded lines, then some of the shed load can be restored. However at this stage, a change of the flow should not cause an overload in any lines in the system. This is possible if the lines with phase-shifting transformers feed the additional power. Suppose line i is connected between buses a and b and has a phase-shifting transformer which is adjacent to bus a. The real power flow f_i on this line is given by,

$$f_i = V_a V_b Y_{ab}\ cos(\theta_{ab} - \delta_a + \delta_b) - V_a^2 Y_{ab}\ cos(\theta_{ab})$$

If $\nu_{ab} = \delta_a - \delta_b$ represent the bus angle increment, then the change in real power flow with respect to the change in the bus angle increment is given by,

$$\Delta f_i = \frac{\partial f_i}{\partial \nu_{ab}} \times \Delta \nu_{ab}$$

Let U_{ia} be the sensitivity function, defined by the following equation,

$$\begin{aligned} U_{ia} &= \frac{\partial f_i}{\partial \nu_{ab}} \\ &= V_a V_b Y_{ab}\ sin(\theta_{ab} - \delta_a + \delta_b) \end{aligned}$$

therefore,

$$\Delta \nu_{ab} = \frac{\Delta f_i}{U_{ia}}$$

A change in a power flow will cause changes in bus angles and corresponding changes in other line flows. If the flow change in line i is Δf_i, the change of angle at bus j is given as,

$$\Delta\delta_j = \xi_{j,ab} \times \Delta f_i$$

where, $\xi_{i,ab}$ is given as,

$$\xi_{j,ab} = \frac{(X_{ja} - X_{jb})x_i}{x_i - (X_{aa} + X_{bb} - 2X_{ba})}$$

where X is the element of the bus-reactance matrix, and x is the line reactance. Hence the adjustment required by the phase-shifter is denoted by $\Delta\gamma$ and given by the following equation,

$$\Delta\gamma = \Delta\delta_a - \Delta\delta_b + \Delta\nu_{ab}$$

using this procedure, we can optimize network flows and minimize the required load shedding.

In order to release bus overvoltages in the network, the following alternatives were considered:

- Adjusting control transformers
- Adjusting reactive power injection to the network

Adjustments of tap settings of control transformers would reroute reactive power flows in a power network, and set bus voltages within permissible limits. The most appropriate control transformer for this job is selected depending on the sensitivity of different bus voltages to tap settings of various transformers in the network. These sensitivities are are available as inputs to the expert system program. If these control transformers are not situated in proper locations in the network, reactive power injections to the system would be adjusted as another alternative for releasing bus voltage violations. These selection processes are also based on the sensitivity of different bus voltages to injections of the reactive power into the network.

SEQUENCE OF OPTIONS FOR A SECURITY ANALYZER

As discussed before, the analyzer would consider a specific sequence of remedial alternatives in the security analysis. These alternatives and the corresponding sequence are given as follows:

- Reroute real power flows to alleviate overloads in transmission lines by adjusting tap-settings of phase-shifting control transformers.
- Adjust real power generations schedule to alleviate overloads in transmission lines.
- Shed loads in the system to alleviate overloads in transmission lines.

- Reroute reactive power flows to remove bus voltage violations in the system by adjusting tap-settings of control transformers.

- Adjust reactive power generations schedule to release bus voltage violations.

RULE BASE FORMULATION

In this section we will discuss the corresponding rules implemented in this approach, and steps which are followed by the analyzer to restore the normal operation of a large scale system. These rules are written in such a way that regardless of the type of disruption, the approach would be localized and the technique would be applicable to any size power system. This section is followed by an example for a 30-bus system.

Rule 1: If the power network has overloaded components in the system, then alleviate overloads on those components using control transformers and via rerouting real power flows.

Rule 2: If the power network has voltage violations in the system then restore the voltage profile of the system using control transformers and via rerouting reactive power flows in the system.

Rule 3: If the overloads in the system are not alleviated by rerouting real power flows, then adjust generation power schedule.

Rule 4: If the overloads in the system are not alleviated by adjusting the generation power schedule then perform load shedding.

Rule 5: If the power network has voltage violations after rerouting of reactive power flows then adjust reactive power injections at various buses in the system.

Rule 6: If the real power flow in a line is more than the capacity of that line, the line is overloaded.

Rule 7: If more than one line is overloaded, then list the lines in a descending order, and consider the line with the maximum overload first, for the rerouting of the power flow.

Rule 8: If a specific overloaded line is selected, then consider the most sensitive generator for adjusting its injection, i.e. for line i select the maximum A_{ij} for all j = 1,NG. The adjusted power flow is related to the power injection by the following equation,

$$\Delta f_i = A_{ij} \, \Delta P_j$$

Rule 9: If sufficient generating power is available at bus j, then consider adjusting the generation at that bus as a control action.

Rule 10: Adjusting the generation at bus j may cause, other lines in the system carry overloads. So, adjust the generation at bus j properly such that it would not cause additional line overloads.

Rule 11: If the control of generation at bus j would release overload on line i, then delete line i from the list of overloaded lines, and determine the modified real power flows in all the existing lines in system.

Rule 12: If for a given line i, the control of generation at bus j is not feasible, then according to the given sensitivity factors, consider the next sensitive generator at bus k for alleviating the overload on line i.

Rule 13: If the available control actions for a given line i, can not release the overload on line i, then consider the next line on the list of overloaded lines for alleviating the overload. Continue this process until flows in overloaded lines have been adjusted as much as possible.

Rule 14: If a line is overloaded, and generators available at nearby buses can not be adjusted sufficiently to release the overload, and there is a phase-shifter located on one end of this line, then change the tap-setting of the phase-shifter according to the given sensitivity U_{ij} such that,

$$\Delta f_i = U_{ij} \, \Delta \delta_j$$

Rule 15: If adjusting the phase-shifter would cause a different flow on line i, then calculate the new line flows throughout the network.

Rule 16: If any bus has more than one overloaded line and those lines have power flows in the same direction, i.e. power is flowing into the bus, or the power is flowing out of the bus, then determine the amount of flow that should be reduced from all the lines connected to that bus which have power flowing in the same direction.

Rule 17: If the bus has more than one line, from which the real power flow should be reduced, then identify the sum of the flow reductions in all the lines connected to that bus for incoming as well as outgoing overloads.

Rule 18: If for a given bus the incoming overload is greater than the outgoing overload, then shed the load on that bus by the amount given as, (incoming overload) - (outgoing overload).

Rule 19: If for a given bus the outgoing overload is greater than the incoming overload, and that bus is a generating bus, then reduce the generation at that bus by the amount given as, (outgoing overload) - (incoming overload).

Rule 20: If load has been shed at specific buses of the system, then make a list of all those buses and arrange them in descending order, starting with the bus which has the maximum load shedding.

Rule 21: If the list of buses with load shedding is non empty, then consider the first bus on the list, and make a list of lines which are feeding power to this bus and have a connection to a phase shifting transformer.

Rule 22: If more than one line is available for restoring the load at a bus, then consider the line with maximum available margin first, and calculate the amount of real power flow adjustment, Δf, as follows,

$$\Delta f = \begin{cases} \mathit{load\ shed}, & \text{if load shed} < \text{line margin} \\ \mathit{line\ margin}, & \text{if line margin} < \text{load shed} \end{cases}$$

Rule 23: If for a given line the amount of adjustment of the real power is known, then calculate the proper tap setting of phase shifting transformer, and determine the revised status of the power system.

Rule 24: If for a given line the amount of adjustment of the real power is known, then calculate the proper change in the generation schedule using sensitivity values which represent the change in real power flow with respect to changes in real power injection.

Rule 25: If the voltage at a given bus is more than the maximum permissible voltage or less then the minimum permissible value, then identify that bus as the one with voltage violation.

Rule 26: If several buses have voltage violations, then consider the one with maximum voltage violation first.

Rule 27: If bus i is selected for adjusting its voltage violation, then consider the D sensitivity factors and identify the most sensitive bus with the reactive power injection for this adjustment.

Rule 28: If the most sensitive generating bus is identified for the adjusting its reactive power injection, then make a proper change in the reactive power injection at that bus and calculate the new voltage magnitudes at all the buses in the network.

Rule 29: If a bus has voltage violation, and there are no nearby generating buses with adequate reactive power injection, and the bus is equipped with a tap changing control transformer, then consider the T sensitivities and adjust the setting of the tap-changer accordingly.

Rule 30: If the proper adjustment of a tap-changer is available at a specific bus, then adjust the setting of that control transformer and calculate the new voltage magnitudes at different network buses.

RESULTS

As discussed earlier, the power system security analyzer uses various methodologies for the power system restoration in an emergency. In order to test the performance of the analyzer, an IEEE-30 bus system, shown in Figure 4, is considered with a given contingency which is studied as follows.

Fault : Lines 1, 4, 5, 6, 12, 21, 24, and 39 are overloaded by 5.0MW, 9.0MW, 6.0MW, 4.0MW, 1.0MW, 3.0MW, 1.0MW, and 1.0MW respectively. Also, buses 26, and 30 have voltage violations of 0.013 p.u. and 0.01 p.u. respectively.

Action : At first, we would consider line overflows. Therefore, phase-shifter transformers on lines 5 and 21 are selected for phase angle adjustments. The phase-shifting transformer on line 5 would be adjusted by 0.54 degree and the one on line 21 is adjusted by 0.32 degree. Since overloads have not been removed completely from the system, generators at buses 2, 5 and 11 are selected for adjusting real power injections. The injection to bus 2 would be decreased by 4MW, injection to bus 5 is increased by 17MW, and injection to bus 11 needs to be decreased by 3MW. Buses 4, 5, 7, 8, 15, 16, 17, 20, 21, 29, and 30 are selected for load shedding by 3.0MW, 2.0MW, 3.0MW, 1.0MW, 3.0MW, 1.0MW, 1.0MW, 2.0MW, 4.0MW, 2.0MW, and 2.0MW respectively. The line flow solutions at this stage indicate that the system has retained its normal state. However, for the optimization process, phase-shifting transformers on lines 3 and 40 are selected to restore a fraction of loads at buses 4 and 30. The phase shifting transformer on line 3 would be adjusted by 0.31 degree and the one on the line 30 is adjusted by 0.95 degree. We would consider bus overvoltages at this stage. So, the reactive power compensator at Bus 27 is selected for adjusting the reactive power injection. The injection at bus 27 would be increased by 5.4MVAR.

Result :

> Overloads on lines 1, 4, 5, 6, 12, 21, 24, and 39 are released. Load flow results for adjusting the generation schedule and load shedding are given in Table 1. Voltage violations on buses 26, and 30 are released, and load flow results for bus voltages once the reactive power injection has been modified are given in Table 2.

So, in an emergency situation where the integrity of a large power network is jeopardized, it is a common practice to reroute power flows through alternate paths or shed non-critical electrical loads so that the least number of customers get affected in terms of their electrical supply. Generally, load shedding is not much recommended due to the loss of revenue to the utility, as well as creating unsatisfied customers. On the other hand, due to economical reasons, present day transmission networks, carry large amounts of power, and rerouting the power flows or adjusting the taps of phase shifters may not be sufficient to alleviate the overload in the system. Hence, it becomes necessary to resort to load shedding as one of the key options in the restoration of a power system. Keeping all these points in mind, one has to develop a scheme that satisfies if not all but as many criteria as possible in the reliable operation of a power network.

MAN-MACHINE INTERFACE

The power system security analyzer is developed on a HP 9000/series 330 workstation, with a HP-UX operating system, using a HP Windows/9000 environment. The graphical representation of a power system status is the most convenient and natural way for system operators to perceive the state of the power system at any moment. Factors involved in a man-machine interface are given in Figure 3. However, the output for the analyzer is in graphic as well as alpha-numeric formats. For this specific application, the analyzer utilizes the HP Windows/9000 (HPW) environment. The HPW environment allows the display of more than one window on a single output device. The analyzer uses three windows on the display device. Out of three windows one is a graphic window named "layout" and the remaining two are alpha-numerics named "expert-sys" and "sys-access".

The graphic window "layout" displays the power system layout in a one line diagram format. Different states of transmission lines are displayed using different colors. Loads, generators, phase-shifting control transformers, and tap-changing control transformers are all displayed using various symbols. The transmission lines are in one of emergency, alert, or normal states. These three states are represented in red, yellow and green colors respectively, which gives an operator a graphic display of loadings on various transmission lines in the system. The other important quantity from the operator's point of view is the actual flow in transmission lines, and to meet this requirement the analyzer displays two numbers in yellow and red color for each transmission line. The number in yellow represents the actual flow and the one in red represents the maximum

Table 1

IEEE 30 Bus Results – Line Overload Alleviation Solution

Line No.	Connection Between Buses	Actual Flow (MW)	Flow After Adjustments (MW)	Line Limit (MW)
1	1 - 2	74.4	52.2	70.0
2	1 - 3	41.4	32.2	45.0
3	2 - 4	27.0	26.0	30.0
4	2 - 5	48.9	35.5	40.0
5	2 - 6	35.0	24.8	30.0
6	3 - 4	38.7	29.7	35.0
7	4 - 6	37.1	31.4	40.0
8	4 - 12	20.2	15.9	25.0
9	5 - 7	6.2	0.1	10.0
10	6 - 7	29.4	20.2	40.0
11	6 - 8	7.5	6.2	10.0
12	6 - 9	11.0	8.9	10.0
13	6 - 10	10.1	8.4	10.0
14	6 - 28	13.8	12.0	15.0
15	8 - 28	2.5	2.2	5.0
16	9 - 10	31.0	25.9	35.0
17	9 - 11	20.0	17.0	20.0
18	10 - 17	3.1	2.2	5.0
19	10 - 20	7.6	6.0	10.0
20	10 - 21	16.4	15.3	25.0
21	10 - 22	8.0	4.8	5.0
22	12 - 13	30.0	30.0	35.0
23	12 - 14	8.6	7.9	10.0
24	12 - 15	20.5	18.0	20.0
25	12 - 16	10.1	9.0	10.0
26	14 - 15	2.5	1.8	5.0
27	15 - 18	7.5	7.1	10.0
28	15 - 23	7.1	7.4	10.0
29	16 - 17	6.0	5.9	10.0
30	18 - 19	4.5	4.1	10.0
31	19 - 20	5.6	5.9	10.0
32	21 - 22	1.7	1.2	5.0
33	22 - 24	7.2	6.0	10.0
34	23 - 24	4.1	4.3	5.0
35	24 - 25	1.2	1.2	5.0
36	25 - 26	4.1	4.1	5.0
37	25 - 27	2.9	2.9	5.0
38	27 - 28	16.2	14.1	20.0
39	27 - 29	6.1	3.8	5.0
40	27 - 30	7.2	7.5	10.0
41	29 - 30	4.0	3.7	5.0

Table 2

IEEE 30 Bus Results – Bus Voltage Violation Solution

Bus No.	Voltage Before (Kv)	Voltage After (Kv)	Voltage Range (Kv)
1	106.0	106.0	106.0-106.0
2	104.5	104.5	104.5-104.5
3	103.2	103.2	96.0-106.0
4	102.5	102.5	96.0-106.0
5	101.0	101.0	101.0-106.0
6	101.6	101.6	96.0-106.0
7	100.6	100.6	96.0-106.0
8	101.0	101.0	101.0-106.0
9	102.6	102.6	96.0-106.0
10	100.3	100.3	96.0-106.0
11	108.0	108.0	108.0-108.0
12	103.4	103.4	96.0-106.0
13	108.0	108.0	108.0-108.0
14	101.6	101.6	96.0-106.0
15	100.8	100.8	96.0-106.0
16	101.3	101.3	96.0-106.0
17	100.0	100.0	96.0-106.0
18	99.2	99.2	96.0-106.0
19	98.7	98.7	96.0-106.0
20	99.0	99.0	96.0-106.0
21	99.0	99.0	96.0-106.0
22	99.0	99.0	96.0-106.0
23	99.0	99.0	96.0-106.0
24	97.6	97.6	96.0-106.0
25	97.5	97.5	96.0-106.0
26	95.7	97.1	97.0-106.0
27	98.4	98.4	96.0-106.0
28	101.0	101.0	96.0-106.0
29	96.3	96.8	96.0-106.0
30	95.0	96.0	96.0-106.0

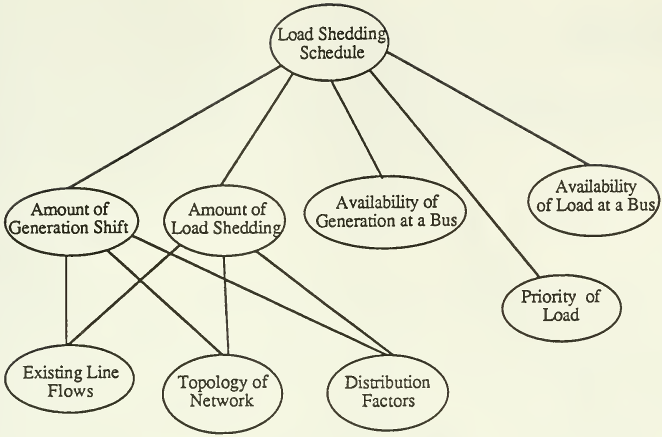


Figure 2 Factors Affecting Load Shedding Schedule

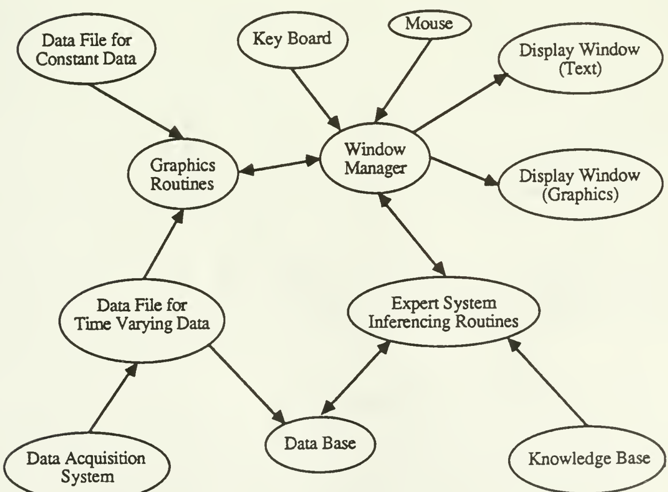


Figure 3 Schematic of Man-Machine Interface

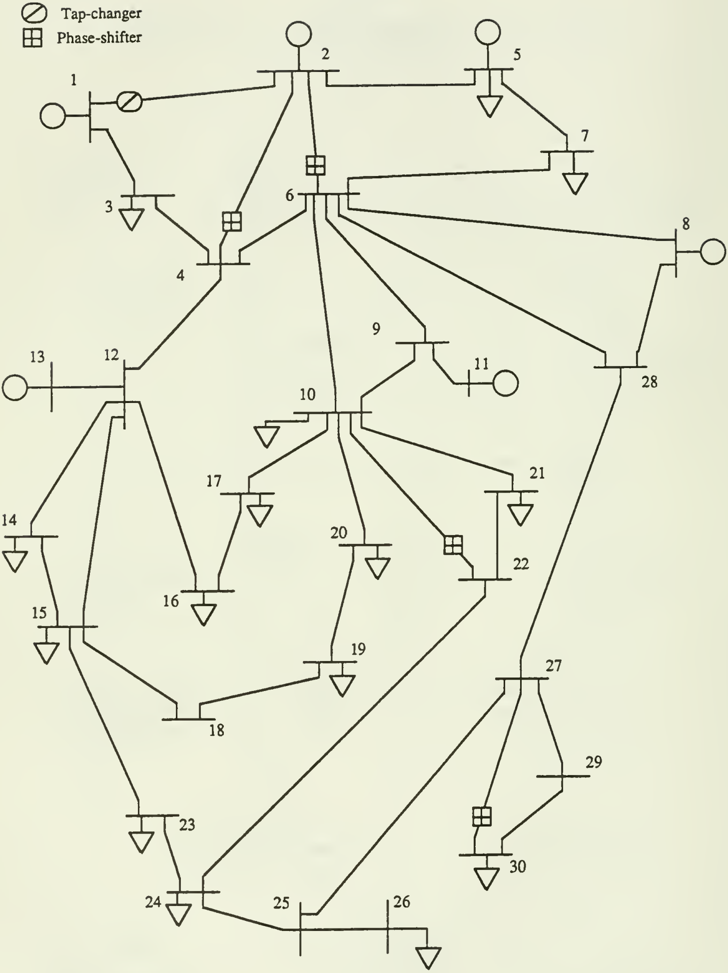


Figure 4 Schematic Diagram for IEEE 30 Bus System

capacity of the transmission line.

The other two alpha-numeric windows, "expert-sys" and "sys-access", are for the communication with the analyzer. The "expert-sys" window gives advice, by printing various control actions, while "sys-access" window gives access to the HP-UX operating system for auxiliary tasks that might be required by an operator. The analyzer makes various suggestions for the corrective action in the "expert-sys" window and then displays graphically the effect of those corrective actions by simulating the post action status of the power system.

CONCLUSIONS

Equipment overloads in a transmission network are caused by unscheduled outages of various components of the network. Since the repair or the replacement of the damaged equipment may require a considerable amount of time, other components which are feeding the loads may have to carry overloads. These overloads may be in great excess of the short-time ratings of these lines. Hence, an operator would have to resort to various options to restore the normal operation of the system. Under such conditions, the system operator is faced with difficulties such as identifying the problem, determining the proper remedial action, and possibly shedding a specific amount of load at right locations. These tasks are difficult to perform particularly if the time is precious.

Generally, a power system security analyzer will act as an aid to the power system operator in making decisions in an emergency situation. In this regard, the status of the power system at any moment, is supplied to the analyzer by the available energy managements system's data acquisition system, thus from the operator's point of view there is not much data that needs to be fed to the analyzer. The development of power system security analyzer, and its validation by testing it on various practical systems gives evidence that the knowledge-based approach is effective in solving power system operation problems which involve highly qualitative reasoning using extensive heuristics. Both qualitative as well as quantitative schemes may be considered, and the transformation of power system data into the symbols and subsequent processing of these symbols may lead to an effective analysis of the power system status. Writing rules to express spatial and temporal context knowledge, and interfacing with the domain expert to refine these rules are much easier in this type of approach compare to the ones which are directly coded in a conventional programming language. The structure used in this study is very flexible, and can be used to solve similar types of problems which involve balancing of load over an interconnected network with several links out of service.

This work has combined the application of many fields of engineering such as knowledge engineering, power engineering, etc., for a real-time application. The power system security analyzer presents a new and viable alternative to minimize the deterioration of the system in an emergency situation that would exist in a power system. A knowledge-based system developed in this fashion would help a power system operator objectify the selection criteria used in power system control which could eventually set standards for the operation of a large power system.

ACKNOWLEDGMENT

The authors would like to thank Mr. M. M. Adibi of IRD Corporation for providing helpful discussions and valuable suggestions. This project was supported by Technology Commercialization Center of the State of Illinois.

REFERENCES

1. S. S. Shah and S. M. Shahidehpour, "A Heuristic Approach to Load Shedding Scheme", Paper No. WP89 - 111, IEEE Tran. on Power Systems, 1989.

2. S. S. Shah and S. M. Shahidehpour, "Automated Reasoning : A New Concept in Power System Security Analysis", International Workshop on Artificial Intelligence for Industrial Applications, Hitachi City, Japan, 1988.

3. S. S. Shah and S. M. Shahidehpour, "Application of Expert System in the Design pf Power System Security Analyzer", Expert Systems Theory ad Application - IASTED International Conference, Geneva, Switzerland, 1988.

4. S. S. Shah and S. M. Shahidehpour, "Application of Expert Systems to Security Analysis in a Power Network Environment", Proceedings of the American Power Conference, Chicago, 1987.

5. S. M. Shahidehpour and G. D. Kraft, "Applications of Artificial Intelligence to Distributed Processing in a Power Systems Environment", Proceedings of the 6th Power Plant Dynamics and Control, Knoxville, Tennessee, 1986.

6. G. D. Kraft and S. M. Shahidehpour, "Recovery Techniques in a Distributed Power Systems Environment", Proceedings of EPRI Seminar on Power Plant Digital Control and Fault-tolerant Microcomputers, Phoenix, Arizona, 1985.

7. J. Qiu and S. M. Shahidehpour, "A New Approach for Minimizing Power Losses and Improving Voltage Profile", IEEE Transactions on Power Systems, Vol PWRS - 2, No. 2, May 1987, pp.287 - 295.

8. John Endrenyi, S. M. Shahidehpour, et.al., “Bulk Power System Reliability Concepts and Applications”, IEEE Transactions on Power Systems, Vol 3, No. 1, February 1988, pp.109 - 117.

9. N. Deeb and S. M. Shahidehpour, “An Efficient Technique for Reactive Power Dispatch Using a Revised Linear Programming Approach”, Electric Power Systems Research Journal, Vol. 15, No. 2, pp.121 - 134.

Development of an Expert System for Electric Distribution Planning and Design

PATRICK M. CAUSGROVE
Paralogix Corporation

RICHARD D. SPERDUTO
Rochester Gas and Electric Corporation
Rochester, New York, USA

DAVID R. WOLCOTT
New York State Energy Research and Development Authority

ABSTRACT

The New York State Energy Research and Development Authority (NYSERDA) and the Rochester Gas and Electric Corporation (RG&E) recognized the need for better planning tools to deal with changing conditions in the distribution of electricity. In response to this need, NYSERDA and RG&E sponsored a development project to create an expert system that aids in solving electric distribution planning and design problems.

The complexity that occurs in planning and designing electric distribution facilities can be managed using the artificial intelligence techniques incorporated in expert systems. In such an expert system, the reasoning mechanisms must work closely with the representation of the distribution plant and take advantage of existing algorithmic methods for analyzing power systems. This intelligent computer-aided engineering system is based on a flexible representation to describe the distribution facilities. An embedded rule-based component interacts with the representation to enable analysis at various levels of abstraction. This processing can be used to reduce computational load or enhance the interactive use of the system.

Planned future developments will extend the capability to encompass distribution operating tasks in the utility.

INTRODUCTION

<u>Background</u>

NYSERDA and RG&E sponsored a project to produce a software system, based on an engineering workstation, which aids distribution engineers in modeling, analyzing, and planning for maintenance, expansion and modernization of distribution circuits. The research development in this project was conducted by Paralogix Corporation, with RG&E acting as the host utility.

NYSERDA recognized the need for better distribution planning tools to deal with changing conditions in the distribution of electricity. Of particular concern was lowering the costs of interconnecting Dispersed Storage and Generation (DSG) facilities to utility distribution networks. NYSERDA considered the application of expert systems as a way to rationalize the process of designing and specifying the connection of these facilities to the distribution network. Concurrently, RG&E saw potential in developing electric distribution expert systems which could be used to enhance distribution reliability, increase operational safety, and improve engineering productivity.

Using artificial intelligence techniques, Paralogix developed the NetReps(tm) network representation scheme which has been the foundation for several expert systems used in computer-aided engineering domains. The LAN/CAD (Local Area Network/Computer Aided Design) system was developed with telecommunications experts for the cable television industry. NYSERDA sponsored a project to adapt LAN/CAD technology to gas distribution engineering design and planning with Niagara Mohawk Power Corporation acting as host utility. GEESE (Gas Engineering Expert System Environment), developed as a result of this effort, has been installed in the RG&E Gas Engineering Department. NYSERDA sought to extend the concepts and the general problem-solving framework developed in these previous systems to the domain of electric distribution.

Distribution system planning, design, and operation at RG&E applies state-of-the-art industry practices. However, the complexity in considering the combination of variables associated with layout, components, cost, and operating performance requires a great deal of engineering manpower or restriction of the variables to reduce problems to a manageable size. RG&E envisioned how the application of the Paralogix technology could contribute to their ongoing efforts to reduce limitations on improved economic management of electric distribution plant assets. The management plan directed that portions of the distribution system be modeled immediately so that RG&E would gain incremental benefit in terms of reduced line loss and more effective loading analysis. Then, as the system became further refined and developed, other distribution areas would be modeled and other application areas implemented. The plan projects application areas to include demand side management, co-generation scenario analysis, and Automated Mapping and Facilities Management (AM/FM) functions.

The Approach

Significant model development costs are required to take advantage of existing algorithmic methods for analyzing real-world utility problems. The computational cost is also very high. The architecture of this inferential computer-aided engineering system is based on a flexible representation to describe the distribution system. Entry of the description is managed by an inferential specification process that can deduce much of the required information and allows descriptive detail to be built in a stepwise manner. A rule-based system component interacts with the representation to allow users to analyze a circuit at various levels of detail. This abstraction, which reduces computational load and enhances interactive use, is dependent on the design or planning context in which the user is working.

Initial development focused on the electric distribution facilities between the substation and the distribution transformer. The work integrates the spatial data representation describing a radial circuit with tools for performing distribution and design engineering analyses on the power system model.

Individuals from the Electric Transmission Distribution and Planning Division at RG&E served as the source of power systems engineering expertise. Several of these individuals are responsible for research and development at both RG&E and at the inter-utility level, thus bringing a high degree of expertise to the project. These people and the staff at Paralogix formed the project development team.

EDaPT

The system that was developed during this research project is called "EDaPT". (Electric Distribution and Planning Tool) EDaPT has two primary elements: Mapping/Data Acquisition, and Planning/Design.

The Mapping/Data Acquisition user of EDaPT enters the distribution circuit drawings into the computer system. A distribution circuit, which usually consists of several maps, is entered, one map at a time, by means of a digitizing tablet. A user can interactively request at any time that a circuit be built from its present set of map sheets - in which case these map sheets are "tied together" at their offsheet reference points to produce a circuit network. Following an incremental strategy, the Planning/Design user does not have to wait until the entire distribution system database is created, but can work with either partial or complete circuit information.

EDaPT graphically displays the distribution system in many levels of detail. Users can view multiple circuits, an individual circuit, or an individual map sheet. These multiple levels of viewing are enhanced by zooming and panning features which allow virtually any portion of the distribution system to be retrieved in a few seconds.

Using graphic displays of the network, the user can interactively modify or query any particular object on the map; e.g. to change a transformer size from 25 KVA to 37.5 KVA, to change the type of conductor, or to determine if a switch is open or closed. Default information is used to reduce data entry by the user.

The coupling of EDaPT's graphical user interface and object-oriented network representation provides a robust environment for developing alternative engineering design scenarios as well as managing the distribution system's data at the operations level.

The user of EDaPT is able to select an area of interest and use engineering tools to analyze it. Users of this component are aided by the Model Builder. The Model Builder employs an integrated inference engine and domain-specific "rules" or heuristics, to reduce complexity while maintaining relevancy of the model for analysis.

Once the Model Builder has produced an appropriate model, the user can submit the model to an analysis subsystem where engineering parameters such as voltage, power flow, and current can be studied on a per-phase basis. The results are displayed using color graphics for quick feedback. EDaPT also provides hard copy results of these analyses.

THE DISTRIBUTION ENGINEERING DOMAIN

The problem definition phase of the project focused on those processes of electrical engineering concerned with planning expansions, maintenance, and modernization of a distribution system. The central goal of the project was to use expert systems to aid distribution engineers and planners in these activities.

Seven major problem-solving areas of distribution engineering at RG&E were identified as summarized below:

1. Correcting operating problems

   The operating departments report problems such as low voltage, frequent outages, or observations related to unbalanced three-phase systems. The distribution engineer must design corrections or enhancements such as circuit reconfiguration, additional use of capacitors, or re-conductoring of lines.

2. Performing sensitivity analyses

   This task is ongoing and performed to predict and prevent problems on portions of the power system that are operating within normal limits but are experiencing load pattern changes. The distribution engineer must design ways to reconfigure the power system and/or design system extensions.

3. Assessing reliability and contingency performance

   The distribution engineer experiments with changing switch configurations in the power system. The engineer must determine for planned or emergency outages if some or all of the load can be picked up by other circuits through switch reconfiguration in the distribution system. This experimentation also gives the engineer information to predict the reliability of the circuit, e.g., identifying single point failures that isolate customers who cannot be picked up by other resources in the distribution system and evaluating their relative exposure to service interruption.

4. Providing for orderly expansion of facilities

   The distribution engineer must design new circuits or extend existing circuits to meet major load additions in a manner that is consistent with the long-range development plan, or planning horizon.

5. Designing changes to distribution circuits in response to shifting load requirements

   The engineer must design system modifications that provide service to the customers, minimize the construction effort, and stay within the planning horizon.

6. Designing system changes for DSG sites

   The distribution engineer must design circuit modifications and an appropriate protection scheme to handle the variable requirements of these sites. It is possible for a DSG site, depending on conditions, to be either a source for power on the circuit or a sink for power, thereby presenting special design considerations.

7. Providing system operational improvements

   Analysis and design activities are required of the distribution engineer to find ways to improve the power system operation by reducing electrical loss, reducing the maintenance cost, or improving the reliability and safety of the system. Knowledge acquisition meetings were held among members of the project development team. Representatives of RG&E's Electric Mapping and Substations Departments were also called upon to lend their expertise. These meetings provided the project with extensive information in the form of maps, standards, and general domain knowledge. Significant time and effort was devoted to understanding prevailing techniques used in modeling and analyzing distribution circuits.

SYSTEM MODEL OF EDaPT

The EDaPT system is an intelligent computer-aided engineering environment that provides the capability to:

1. Obtain a description of the existing power system for a geographic area of interest.

2. Specify alternative circuit configurations as well as constraints, restrictions and evaluation criteria.

3. Model the proposed circuit configurations.

4. Study the circuit models with analytical, heuristic, and symbolic tools.

5. Make decisions based on the resultant analyses.

As a tool for synthesis, the system provides a powerful set of interactive tools to allow complete or incomplete descriptions of distribution circuits. After a circuit schematic has been entered into the computer, the system retrieves valid choices for specification of graphical objects appearing on the display. Objects that are incomplete in specification are given default values by the system, based on object-oriented relationships. In this manner, a working description of the power system can quickly be created. Specific changes and refinements can be made to the rough description to add detail where the engineer desires.

As a tool for analysis, the Model Builder, employing an embedded rule-based system, provides the engineering intelligence to model the distribution system. This procedural knowledge is stored in the Model Reduction Rule Base. The rule base (knowledge base) uses IF-THEN rules (productions) that collectively describe how to transform the distribution circuit(s) into a model suitable for mathematical analysis, i.e. Loadflow Analysis. These rules embody the expertise to reduce detail where not required, yet enhance detail which is important to analysis. For example, the following rule describes the state in which the Model Builder would reduce complexity by "eliminating" a "non-significant" tap. A tap is defined as a branch feeder having a terminal endpoint which is not a switch.

```
IF
      the tap is near the substation
      or
      the tap length is reasonably short
      and
      the tap load is fairly low
      and
      the conductor size is adequate
THEN
      collapse all the tap load to the tap point
```

The Model Builder uses collections of such rules along with forward-chaining inference to synthesize a mathematical model of the circuit(s) to be studied. These rules are coupled with procedures that compute factors associated with the vague terms in the conditional statements, i.e., "probable facts". The rule base EDaPT uses to describe these transformations can be considered independent from the rest of the system, thus serving as a tool for knowledge engineering and lending great extensibility to the architecture.

The analysis subsystem provides the methods by which the distribution model can be studied by standard power systems analysis techniques. The Loadflow Analysis tool, once applied to the model, yields system voltage, current, and line flow values. These values are displayed to the user on the color graphics monitor with a color-keyed information table. Thus, a voltage profile of the system, for any particular phase, can be conveyed quickly to the user. Hard copies of the analytical calculations can also be requested so that the distribution engineer may take printed reports of the system performance from his or her computing session.

SYSTEM FRAMEWORK of EDaPT

A general mapping of the EDaPT System Model on the system framework is shown in Figure 1, System Framework of EDaPT. The following discusses this framework and highlights important development strategies.

System Strategy

In considering the many possible hardware configurations for this project, four basic conditions were considered prerequisite:

1. The hardware must support an open systems architecture, industry standards, and the application development tools described below. By using an open systems architecture, or open computing, developers can select the best support tools and languages for knowledge engineering, software engineering, CAE, and graphics from many software vendors, and developed products can be conveniently ported to other hardware bases.

2. The system should have both significant processing speed and a large memory capacity to adequately support the processing of large, highly-detailed distribution circuits and the heavy emphasis on computer graphics.

3. The computer system must be general purpose in design in that the hardware must support symbolic as well as numeric computing.

4. The system should provide support for engineering workstations and mainframe systems as well as provide the capability for remote terminal access.

These four prerequisites indicated that a high-performance engineering workstation would be best suited for the delivery system. EDaPT is based on a 32-bit engineering workstation supporting the UNIX(tm) operating system, common languages, networking standards, and graphics standards. With such a configuration, EDaPT would be portable across many hardware vendors. The development and delivery hardware that was selected for the project was the Sun Microsystems, Inc. Sun 4/260, a high performance workstation rated at 10 million instructions per second (MIPS). The system has 8 megabytes of main memory, 327 megabytes of disk storage and a 19" high resolution color monitor. EDaPT has been ported to Hewlett-Packard 9000 series systems and can also be delivered on these machines.

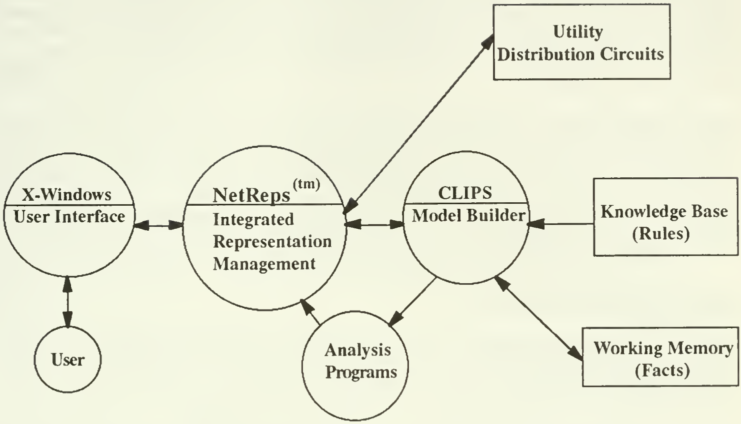


Figure 1. System Framework of EDaPT

Application Development Languages.

The majority of the software used to implement EDaPT is written in the "C" programming language, chosen for its versatility and efficiency. FORTRAN-77 is used to support a loadflow analysis program that was developed by the Energy Systems Research Center of the University of Texas at Arlington. Integrating this program instead of developing this functionality under the project is another expression of the system design strategy to use standards and proven technology within the development. This program is the heart of EDaPT's analysis subsystem, the remainder of which is written in "C".

Network representations and many of the chief data structures used in EDaPT are implemented in NetReps, a proprietary network representation scheme developed by Paralogix, which is written in "C". NetReps has proven to be a useful representation tool in network applications because of its capability to represent and transform different kinds of information in different ways. For example, we not only want to be able to ask our computers questions which pertain simply to counting ("how many things") but also questions which pertain to relationships ("how do these things relate to each other and to utility operations?").

Expert System Development Tools

CLIPS ("C" Language Integrated Production System) was chosen to support the rule-based knowledge representation tools used in EDaPT. CLIPS has many advantages over other expert system "shells". These advantages include:

1. Ease of integration within the UNIX/C environment

   CLIPS was designed to address the delivery problems of integrating and embedding expert systems into conventional environments.

2. Proven track record

   CLIPS was developed by NASA/Johnson Space Center for use in many of their expert systems.

3. Low cost

   CLIPS is available from NASA COSMIC software distribution channels.

Windows/Graphics Environment

The X-Windows system was chosen as the graphics development tool for the user interface. X-Windows, developed at the Massachusetts Institute of Technology, allows the generation of a machine independent graphical user interface. It accomplishes this through a graphics server. This server translates standard requests into the hardware-specific instructions to execute such high level ideas as moving windows on the screen or gathering user input through the keyboard and mouse. The use of X-windows means that none of the user interface routines need be re-written for EDaPT to run on various vendors' hardware.

Data Management

The system framework provides for management of data through standard Unix file system support or relational database management systems. The Ingres database management system was chosen to provide the optional relational database support. Ingres is widely used in UNIX-based software systems, and interfaces well with the "C" language.

COMPUTATIONAL SPECIFICATION of EDaPT

An intelligent computer-aided engineering system was proposed to define a problem-solving environment suitable for the major tasks involved in distribution engineering. This high-level description and the System Block Diagram, Figure 2, present the four major development areas.

Integrated Representation

Development in this area focused on producing software to allow a user to describe an existing power system. The problem of representing the power system in the computer was addressed. The data collection capabilities meet the following specifications:

1. The map and data collection tools must be easy to use.

2. The system must operate normally regardless of whether the power system is fully represented in the computer, or some details are missing.

3. Defaults and inference be widely used to allow quick creation of a rough description of the power system. Specific changes and refinements can be made to the rough description to add detail where the engineer desires.

The integrated representation couples the underlying representational schemes and procedures that are concerned with:

1. Spatial aspects of circuit maps

2. Characteristics and default values for electrical components

3. Methods for traversing/searching the electrical network

4. "Rules of thumb" for reducing the vast quantity of data present in each circuit to an electrical model suitable for analysis

Interactive Modification

This software supports user interaction with the graphics representation of the power system. The software allows the engineer to reconfigure existing circuits, specify the layout of proposed circuit changes, specify new circuits, specify information about circuits, and inquire about circuits and components in the power system.

Special attention was required in this area concerning the routines and services required to implement the man-machine interface for this highly interactive application. Users are given a high degree of control over the workspace on the screen. Windows can be moved around on the screen for optimal placement in relation

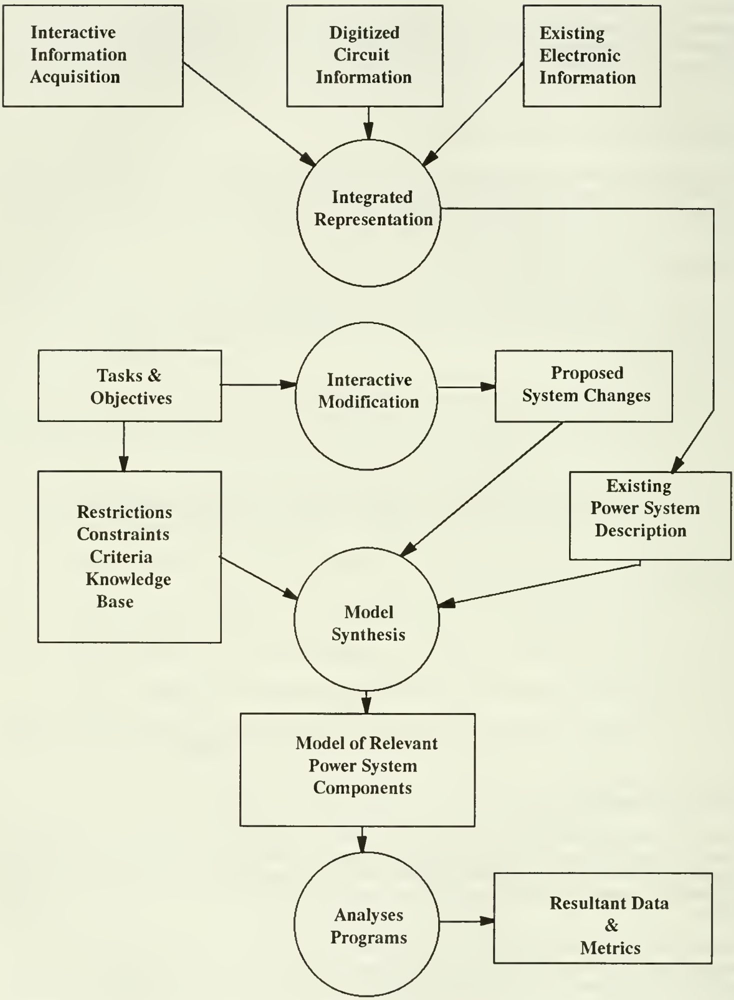


Figure 2. System Block Diagram

to irregularly shaped circuit networks that are displayed. The design requires a minimum amount of typing as most choices are made by selecting graphic objects on the screen with the mouse. Circuit information is shown in a graphical representation using shape and color to signify meaning, allowing users to interpret the data much more quickly than by examining tabular reports.

The graphics interface is closely coupled with the underlying integrated representation to encode not only the graphic elements of an object but also the meaning of the object to the expert system. Thus, the graphics information becomes a valuable component of the overall cognitive activity of the system.

Model Synthesis

This software takes the physical description of the system and transforms this description to a data structure suitable for mathematical modeling of the power system. The transformation considers at least five factors:

1. The kind of analysis to be run
2. The problem the analysis is intended to help solve
3. Planning criteria
4. Design constraints
5. Common practice in model definition

The reasoning mechanism, knowledge framework, and computational specification implemented in this software are general in nature. This implementation provides the capability to perform the Model Synthesis task and can be extended easily to handle design synthesis, application of planning expertise to create layout and operating plans. It was observed that much of the thinking that is applied to create an appropriate and compact model for a planning scenario is similar to the thinking involved to select and lay out a solution in circuit design. An incremental approach was taken which provides a general foundation for building new expert behavior in response to additional requirements.

Analyses Program

The system employs standard mathematical methods to analyze distribution system performance in terms of power flow calculations and voltage profile. This subsystem provides the algorithms and mathematical techniques used in power system analysis, such as the Newton-Raphson iterative power flow.

CONCLUSIONS

Applications

This computer-aided engineering tool is beneficial in allowing users to simulate the effect of proposed changes to the distribution system between the substation and distribution transformers. EDaPT provides a utility with the means of creating a database of distribution facilities incrementally in response to operating needs. The engineer is no longer required to adapt a distribution circuit model for different analyses of the same geographic area; EDaPT quickly creates a new model to suit the problem. Yet, even as modeling activity increases, EDaPT ensures consistency between separate planning evaluations which allows a utility to define standardized planning strategies. In addition, once these facilities are described, this information can be applied to benefit other areas of the corporation.

Benefits

An electric utility using EDaPT gains numerous benefits. These benefits include:

1. An increase in productivity and reliability. Engineers are able to propose and evaluate design scenarios some 10-20 times faster than with conventional methods. Conventional methods require from several hours to several days to derive a model and analyze it. EDaPT can produce a model of the system and analyze the model in a few minutes. In addition to the time savings, color graphics are a more effective means of interpreting results.

2. Source data is readily available in the form of the utilities' primary maps. Once captured, circuit information is easily accessed and used to solve a variety of problems.

3. The distribution database can be built incrementally with payback at each step.

4. Newly hired staff learns faster using an integrated tool with domain knowledge.

5. EDaPT is extensible and can also be used to manage data at several operational levels, thus reducing the amount of information recorded manually and enhancing the availability and dissemination of the data sources.

6. EDaPT is not bound to any particular hardware vendor and can run on many different hardware configurations.

7. RG&E employees involved with the development and use of EDaPT have assigned a high value to the degree of control and opportunity presented by the localized databases of the kind in EDaPT. They now can create, maintain, and use this information directly from their own desktops. However, the distributed aspect of the system framework provides communication and connection that makes the data widely available for other corporate uses. Within this type of framework, additional computing horsepower and memory can be added over time to create access to the local database as its corporate value increases, yet be done in a fashion that provides data security.

Electric utilities are always seeking better and faster ways to model their circuits and manage their facilities. This research, by addressing these problems, indicates that a commercial product offspring of EDaPT is likely to succeed in the utility marketplace.

Planned Development

Future development will encompass a broader set of utility planning and operating functions by applying the system framework to extend the knowledge and capabilities. Knowledge acquisition relating to optimized distribution circuit layout was performed in parallel with software development during the project. A knowledge system applied to this problem must establish criteria for collectively evaluating reliability, voltage profile, losses, and capital costs. The decision-making must also take into account the need for the proposed circuit design to consider the long-range planning horizon for the distribution area. This design synthesis system will allow a utility to easily quantify numerous expansion scenarios while documenting the assumptions and constraints considered.

The delivery system and EDaPT are installed at RG&E and are being used to help solve problems. Meanwhile, additional applications are being developed through the ongoing efforts of RG&E and Paralogix. Figure 3, Application Areas, illustrates the numerous directions that can be taken to capitalize on an integrated, flexible representation of distribution facilities. Based on the strength of the use and benefits of the system, RG&E and Paralogix are working to create a commercially packaged implementation of EDaPT.

ACKNOWLEDGEMENTS

The authors wish to formally thank NYSERDA, RG&E, and Paralogix for funding this project. The project development team included Thomas E. Dasson, Thomas P. Frantz, Robert H. Jones, and Carl H. Warn from RG&E and Robert J. DelZoppo, William A. Gimbel, and Edward D. Rafalko from Paralogix. We greatly appreciate the work of the project team and are fortunate to showcase their achievements.

Other individuals at RG&E who were essential in the success of the project were William A. Brewer, Kenneth R. Mullen, Michael B. Whitcraft, Howard Worden and Bernard W. Zapf. Thank you. Dr. Raymond R. Shoults from the University of Texas at Arlington was a considerable help in integrating loadflow analysis, we applaud his collegial attitude.

TRADEMARKS AND NOTES

EDaPT trademark request is in progress.
INGRES is a trademark of Relational Technology, Inc.
NetReps is a trademark of Paralogix Corporation.
UNIX is a trademark of AT&T.

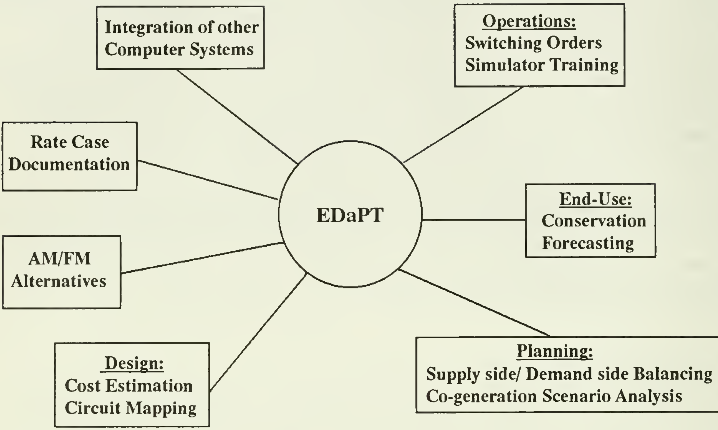


Figure 3. Application Areas

On-Line Condition Monitoring of Power Station Components Using Expert Systems

G. LINDBERG and P. JAUHIAINEN
ASEA Brown Boveri AG
CH 5400 Baden, Switzerland

ABSTRACT

In today's power industry there is a strong tendency to reduce production costs. This goal can mainly be achieved with condition-based maintenance and optimal process control.
Although many power plants do have an extensive and complete instrumentation set-up, this vast amount of information is not normally systematically followed up, analysed and stored. In many cases the operators receive no significant information before alarm and/or trip levels are reached. The Condition Monitoring System, now under development within the authors' company (ABB), is intended to improve the present incomplete systems. With a computerized analysis of trends (e.g. bearing temperature or generator winding temperature) small changes in component behaviour can also be detected. To be able to systematically analyse the deviations of the large amount of signals, Expert Systems have been integrated into the Monitoring concept. By dividing the power plant into a number of components or functional groups, different modules are developed, each comprising its own knowledge base.

As a result of the modular approach the Condition Monitoring System is flexible and can be tailored to the specific needs of a particular power plant configuration. To maintain a high degree of standardisation, the system is implemented and delivered on a VAX-computer.

The aim of this paper is to give the background of and the need for such systems. Furthermore, the system function is described and in particular the use and the implementation of Expert Systems are emphasized.

WHY CONDITION MONITORING?

Nowadays utilities worldwide show a strong interest in the use of Condition Monitoring although the reasons for this may differ from country to country. There appears to be a relationship between the interest of the Management in introducing Condition Monitoring and the educational level and experience of the power plant staff. The Management believes that the introduction of the knowledge-based Expert System increases the independence from the specially skilled personnel. This, however, is only valid to a certain extent. The aim of the Expert System is not to take over the role of the specialist but to support him/her in his/hers work.

In Europe, for example, it is becoming more and more difficult to build new plants because of government regulations, so that the need to extend the lifetime of existing plants increases. The introduction of advanced On-Line Condition Monitoring enables the early detection of changes in the thermal and mechanical conditions of the plant which may otherwise cause a malfunction or severe breakdown of the plant.

Another trend which has been noticed in Europe and the United States for some time is the interest of the insurance companies in encouraging utilities to install Monitoring systems. As the installation of such systems decreases the risk of damage, the insurance fees can be reduced and the power plant owner can achieve a quicker return of the investment.

GENERAL SYSTEM PHILOSOPHY

Before starting the project a feasibility study was made to determine the customers' needs and ideas. When compiling the suggestions of the utilities, a number of fundamental features became evident:

- The system should cover the whole plant.
- The system should be directly accessible and available 24 hours a day.
- The system must be flexible and allow the input of new knowledge.

As a complete set of knowledge cannot be stored in the Expert System, it may be necessary to contact the manufacturer in some cases after a diagnosis has been made. It is unlikely, however, that more than a minor number of actions of the system will include a recommendation to contact the manufacturer.

It soon became very clear that a more powerful and versatile (e.g. multi-tasking) computer architecture was needed to fulfill the functional demands of the system. A VAX-computer (VAXstation 2000) was therefore chosen, using the VMS operating system. With this solution ABB has a hardware concept which is available worldwide and which complies with ABB specifications.

The main goal of the ABB Condition Monitoring concept is to increase the economic efficiency of the power plant. Firstly, the early detection of damages shall prevent consequential damages or at least reduce them. A condition-based overhaul planning increases the availability of the plant by reducing forced outages, see Fig. 1. Secondly, the heat rate or thermal efficiency of the plant can be improved by assisting the operating personnel in an optimal control of the process.

An example is the change in the condenser vacuum due to a deterioration of the tube bundles in a nuclear power plant. This parameter is of much greater importance in nuclear stations than in fossil fired stations because of the relatively short steam expansion line.

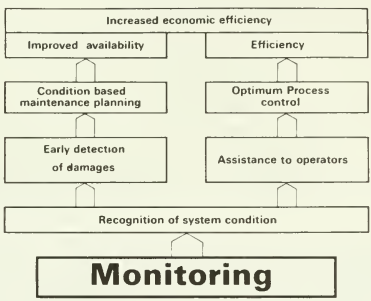


Fig. 1: Main goals of Condition Monitoring

It must be possible to implement and use the Monitoring system in power plants with data acquisition systems of different capability and degree of modernization. Older plants have fixed wires from the sensors to the gauges in the control room whereas modern plants have computerized control systems with data highways. In order to be flexible, ABB has chosen a standard interface, based on VAX standard Ethernet (IEEE 802.3), between the VAX computer and the process control system, see Fig. 2. It is planned to equip older plants, which have no bus system, with a variant of a computerized control system "PROCONTROL P" (ABB control system), which will be connected to the VAX computer by a coupler. In new plants, where ABB PROCONTROL P is already installed, only the data communication interface (coupler) needs to be fitted. In power stations with a non-ABB bus-based control system the coupler must be adapted to the existing control system. A connection to the ABB MASTER control system (all interfaces based on Ethernet) can also be provided.

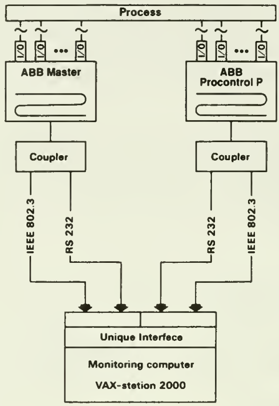


Fig. 2: Connection of the Monitoring system to the process

Before evaluating the data, the system determines the mode of operation (main mode and sub-mode of operation).

The Monitoring system is designed to give additional support to the operator. As the system is completely passive, there is no interaction with the safety system of the plant.

To fulfill the varying requirements of the customers, the Monitoring system is designed as a modular system which permits selection of one of the modules, or even segments of a module, or the entire system, see Fig. 3.

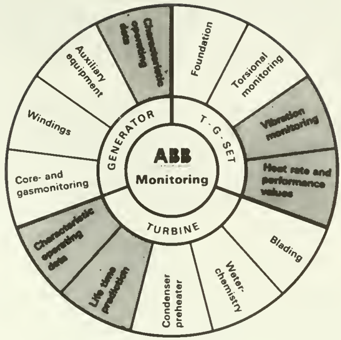


Fig. 3: Modules of the Monitoring system

At present, the following modules are being developed:

- Module 1: Characteristic data of the generator
- Module 2: Characteristic data of the turbine
- Module 3: Lifetime prediction
- Module 4: Heat rate and performance values
- Module 5: Vibration monitoring

During normal operation the system is passive for the operator. If one of the significant parameters, which are monitored (e.g. bearing metal temperature), reaches the warning level, the system reacts. If requested, a diagnosis is given and adequate actions are proposed. This pattern, however, is not adhered to followed by the module "Lifetime prediction", which makes no diagnosis but indicates the remaining lifetime of the examined parts, based on the number of cycles and operating hours.

FUNCTIONAL DESCRIPTION

The On-Line Condition Monitoring system assists the control room operators. The system is passive and does not interact with the normal safety system of the power plant.

In normal operation when warning levels are not reached all internal functions such as the data acquisition, evaluation of process performance values and storage etc. run in the background mode. In case of abnormal conditions, indicated by one of the modules, the operator can make further investigations with a menu-controlled system. This philosophy is used particularly in the module "Vibration monitoring" where the user has a wide spectrum of user-controlled menus and windows for additional analyses (integrated in the front end, TVM-50 or TVM-300).

The Monitoring computer is connected to the power plant control system by the coupler, see Fig. 2. The process data (temperatures, pressures, differential pressures, displacements etc.) are transferred from the control system to the data storage buffer (process image, PI) of the VAX. A new update is made every 10 seconds (maximum: 1000 values), see Fig. 4.

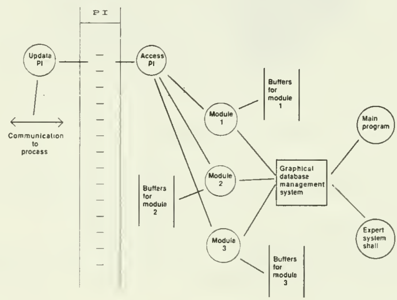


Fig. 4: Internal data flow and storage philosophy of the Monitoring computer

Based on this PI, every module will update the specific module buffers at a frequency which depends on the module. For modules covering only the steady-state condition special routines, such as mean value calculation over time, are planned before the measured values are used for calculation, storage and display.

The process control system checks all measured data for irregularities, and the status check of the measured data is given for every value transmitted to the PI. The next step is a plausibility check, using physical facts, for example:

- In a feedwater line operating normally the feedwater temperature must increase upstream.

- In a steam pipe with a two-phase flow condition the measured temperature cannot be higher than the saturated steam temperature corresponding to the existing pressure.

The mode of turbine operation is determined from the measured values, see Fig. 5.

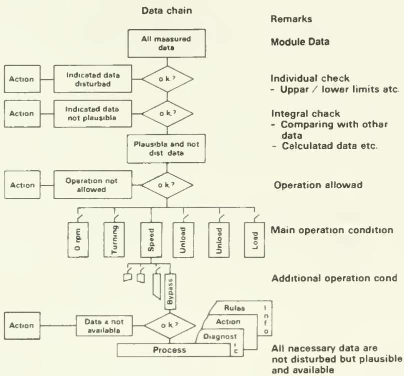


Fig. 5: Plausibility check and mode of operation

The main modes of operation are:

- No rotation of the rotor, 0 rpm
- Turning gear in operation
- Speed operation
- Full speed, breaker open
- Full speed, breaker closed
- Load operation.

In addition to the main modes, submodes of operation are also defined. For the main mode "Load operation", for example, the submodes are the following:

- Load increase
- Steady state operation (with given criteria)
- Load reduction.

Only after establishing the mode of operation, the diagnosis/evaluation can be continued.

In normal operation, the system will only show the main menu and indicate if any of the modules has issued one or more warning alarms. If so, the corresponding module is indicated on the screen. In order to confirm the indication, the user must acknowledge the alarm and can choose between diagnosis (using the Expert System) and evaluation (e.g. trend analysis). After acknowledging the alarm, the module returns to a non-active mode. The alarms are stored on an alarm list which can be shown or printed out on request.

In case of an alarm, the user has three possibilities:

- Evaluation
- Diagnosis
- Cancellation of the alarms.

In the user-controlled evaluation mode, the procedure to follow is indicated in the menu. The user may wish, for example, to have a trend analysis on the basis of the warning alarm parameters.

As a rule, the parameters also contain information before the alarm levels are reached. The protection functions usually comprise a trip level and an alarm level. This means, however, that the operator does not receive any information on the trend of the measured values before the alarm level is reached. The measured values therefore include many data which are not presented to the operator.

The Monitoring system, however, processes the information of the measured data before the protection alarm level is reached. This function is achieved by introducing an additional warning level below the protection alarm level. Upon the user's request, the warning level response can initiate a trend analysis which permits prediction of the time elapse up to a protection alarm. The time elapsing before tripping is predicted in a similar way, see Fig. 6.

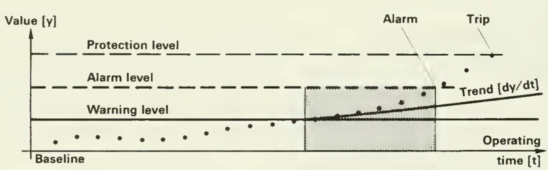


Fig. 6: Trend analysis including prediction using the warning level

The specific modular calculation comprises the evaluation of changes in the thermal or mechanical condition of the plant. These evaluations are always made in the background mode at intervals, depending on the module.

The interpretation of the isentropic efficiency of an IP turbine is given as an example (module "Heat rate"), see Fig. 7.

At intervals of 6 minutes, the actual value of the isentropic efficiency is calculated from the measured and averaged temperatures and pressures at the steam inlet and outlet of the IP turbine. The target value of the isentropic efficiency is also calculated using other measured values such as the load. The values are compared and the difference between actual and target value is used as input to the Expert System. In the user-controlled evaluation mode the operator can also obtain a trend analysis of the isentropic efficiency.

Taking into account the change in the isentropic efficiency and other measured process data such as the swallowing capacity of the turbine, the chemical quality of the feedwater etc., the Expert System delivers a diagnosis of the possible causes and recommends remedial actions.

There are two data storages in the specific module buffers:

- Short-term storage up to 24 hours
- Long-term storage.

All data which are relevant for the diagnosis and/or evaluation are stored in the short-term storage whereas the long-term storage contains only significant data.

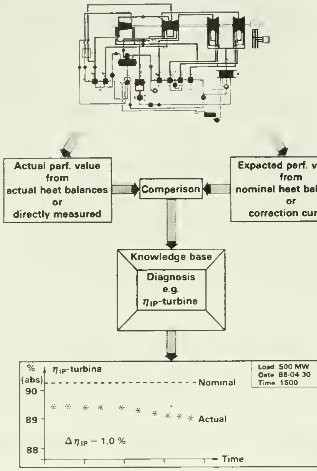


Fig. 7: Outline of the interpretation of changes in the IP isentropic efficiency

APPLICATION OF THE EXPERT SYSTEM

The research activities in the field of Artificial Intelligence (AI) to approximate human behaviour with computer programs has covered fields like natural language understanding, speech, planning systems, robotics and Expert Systems.

An Expert System is a computer programme which is able to solve a given problem within a well-defined and restricted problem area using knowledge represented in the computer to approximate the behaviour and ability of a human expert.

Many Expert Systems have been developed in different areas, most of them as advisory or diagnostic systems [1]. It is important to remember that the Expert System applications are not systems which replace human experts but support the user efficiently and fast in this problem-solving activity.

A description of an expert system can be divided into three parts,

- the knowledge base
- the inference mechanism
- the application interfaces.

The knowledge base contains the information of the specific problem area in which the Expert System application is developed. The information is structured and stored to represent the knowledge of human specialists. The information can be represented in different ways, the most common representation models are rules and objects, others are frames, semantic nets, procedural languages and logical expressions. Many problem areas are not suitable for being represented in a single representation model due to the resulting complexity. These need multiple representation models which are also provided by many Expert System shells.

The inference mechanism is a mechanism that uses the information in the knowledge bases to draw conclusions in order to solve the application-specific problems [2]. The main tasks of the inference mechanism are

- to check which facts in the knowledge base are relevant to the specific problem to be solved and draw conclusions from the results, if possible

- to specify the order in which the search for the facts is to take place.

The explicit separation of representation and inference is the distinctive feature of knowledge-based systems. As a result of this distinction, it is possible to change or extend the knowledge base without changing the inference machine. Compared with other conventional computer information systems, this ensures essentially shorter system development times and also helps to maintain and modify the application, depending on future demands.

The application interfaces are all the interfaces needed for a complete software system. As the Expert System is only a subsystem of the Monitoring system, it is necessary to define the interfaces to the

- data acquisition system

- external calculation programmes (which can be written in other languages than the tool itself)

- end-user graphics.

The Expert System Part in the ABB Condition Monitoring Project

The Expert System in the ABB Condition Monitoring project is a diagnostic tool. It gives a diagnosis of the possible causes of deviations of the measured data in the power plant and recommends corrective actions to the user. The modules also have specific requirements regarding the plausibility checks of the measured data and operating state of the plant. These additional requirements are covered by the Expert System.

The Expert System in no way controls or influences the processes in the plant or its components, it merely recommends corrective actions to the user.

A diagnosis can be made when the system detects measured data deviations which exceed the permitted values. The detection of any deviation is called an "event".

The results of a diagnosis are

- the description of the event

- an explanation of the event

- a certainty factor to indicate the probability according to system knowledge

- recommendations for actions to avoid subsequent damage to the plant.

In every Expert System application the most difficult problem to be solved is knowledge acquisition. Each of the modules in the Monitoring system is usually developed by two specialists. Their experience gained in many years of field-service, e.g commissioning and trouble-shooting, and the knowledge obtained from handbooks and other literature on module-specific problems are the basic input of the module. This draft material is then refined by the knowledge engineer in a form suitable for being implemented in prototypes. The prototypes are further developed to provide the final knowledge bases in the Monitoring system. The knowledge require has so far been acquired by the knowledge engineer, but the final aim is to have it done by the specialists themselves.

The module "Characteristic data of the turbine", for example, uses a commercial object-oriented rule-based shell as an Expert System shell. The knowledge is represented in rules in the logical format

IF (premise) THEN (conclusion) DO (action)

This means that if the conditions of the "premise" are valid, the "conclusions" are also valid and any possible "action" will be carried out.

The different modules have their own knowledge bases where the rules are given different priorities so that the rules concerning more essential information are applied first. This means that the diagnosis is directed to the rules where the probable causes of the event can be found. The specialists apply the same method during trouble-shooting in order to find the cause of a failure.

To confirm and better understand the conclusions drawn from the diagnosis, it is important to give a detailed explanation of the reasons and conclusions for a specific diagnosis of the system. The explanation is an application-specific part which is performed in an external program and is not supported by the Expert System shell.

The certainty factors weight the reasons for the diagnosis according to the system knowledge, i.e. a high certainty factor shows that the diagnosis is well supported by the system knowledge whereas a low certainty factor indicates that there are only certain indications in the system knowledge which support the diagnosis.

As mentioned before, the Expert System is a subsystem of the Monitoring system which must communicate with other software packages. Both the input and output of the Expert System must be defined. The input is the data acquisition system which continuously feeds the values measured in the power plant into the knowledge bases. The output is the end-user graphics which is most important for the end-user acceptance of the system. The data acquisition system and the end-user graphics are external programs of the Expert System.

A schematic diagram of the data and knowledge flow in the Monitoring system is given in Fig. 8.

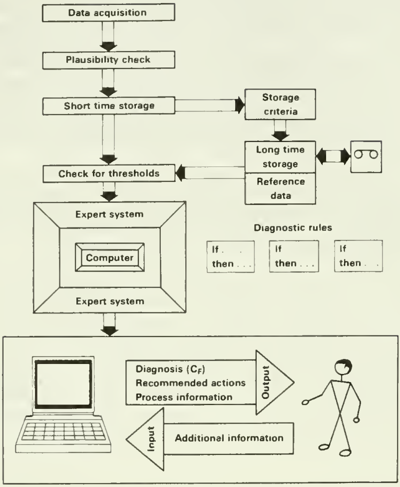


Fig. 8: Schematic diagram of the data and knowledge flow

THE CUSTOMIZED AND MODULAR APPROACH

To meet the customer's demand for flexibility, the Monitoring system is subdivided into a number of modules as shown in Fig. 9.

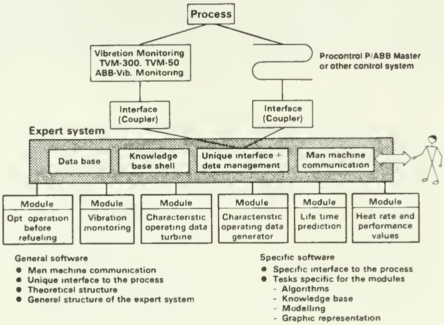


Fig. 9: Modular design of the Monitoring system

On the customer's request the Monitoring system can be supplied in two steps:

Step 1: Monitoring system excluding the Expert System

Step 2: Additional Expert System Part

This means that the customer can start with a less expensive solution and still have all the evaluation facilities at his disposal. At a later stage he can add the Expert System.

Module "Vibration monitoring"

In present-day power plants, vibration monitoring is limited to the indication and recording of the vibration amplitudes. If one of the predetermined limit values is exceeded, an alarm is given and/or the turboset is tripped. This ensures the minimum protection of the plant.

The development of the modern functional TVM-50 and TVM-300 Vibration Monitoring Systems was based on the experience gained with the commissioning and maintenance of turbosets. The systems comprise comprehensive signal conversion and processing which are required for the advanced analysis of the vibration curves obtained from the plant equipment. Using the FFT analysis the measured vibration signals are processed and the results displayed to the user in a variety of diagrams. The system is designed in particular to observe and record the vibrational behaviour during startup. The result, to be called up at any time, can either be displayed on a screen or printed out, see Fig. 10.

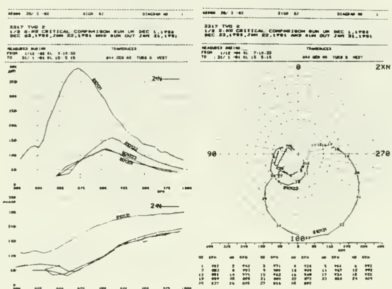


Fig. 10: Runup diagram

The Vibration Monitoring system can be used as a "stand-alone" system or be combined with an Expert System (Fig. 9). The system automatically recognizes the alarms which are checked against the reference values. In case of deviations, the Expert System is started upon request and a diagnosis with adequate actions given. In order to be able to take into account other relevant data, the condenser vacuum, bearing metal temperature etc. are also measured.

The Vibration Monitoring unit is of a compact design and can be integrated into the control room without difficulty. The vibration sensors in existing operating turbosets can usually be connected to the monitoring unit, regardless of whether they are of the relative or absolute type.

Module "Characteristic Operating Data of the Turbine"

In modern power plants the most important parameters are usually measured by continuous line recorders. These values include:

- Overall turboset data (electrical power output, voltage, current, rotor speed, vibration amplitudes, differential expansion, eccentricity, valve positions, etc.)

- Bearing data (metal temperatures, lubricating oil temperature and pressure)

- Metal temperatures (HP and IP turbine casings, valves, pipes, etc.)

- Thermodynamic data (live steam temperature and pressure, wheel chamber pressure, exhaust pressure, etc.)

- Mass flow of the condensing and feedheating equipment.

If these parameters are taken separately, it may be difficult to detect any malfunctioning. If, however, a combination of these parameters is considered, a fault can be discovered earlier. The ABB approach is to compile the measured data in functional groups with only a minor relationship between the groups or no relationship at all, see Fig. 11.

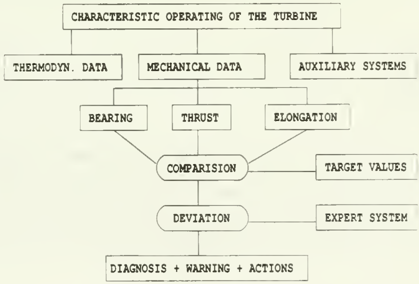


Fig. 11: Brief description of the module "Characteristic data of the turbine"

Based on the deviations resulting from direct measurements or observations and validity tests, a number of fault hypotheses can be established and their probability determined. The measurements of the physical properties are the basis for any assessment. To permit checking of the measured values, special plausibility rules are set up by integrating other parameters with a physical interrelation. When determining the difference between the measured and the expected value, the expected value is always referred to a specific mode of operation. The target values are usually determined by a quadratic function with the load as main parameter.

The deviation of a measured value, for example of the bearing metal temperature, is evaluated in the Expert System which gives a diagnosis and a probability for the possible hazard. If the diagnosis indicates an abnormal condition (with some degree of probability), the system issues a warning and recommends corrective actions. The recommendations can include:

- gathering further information by mobile or local instrumentation

- operating the Expert System with other parameter variations in order to increase the probability of a given diagnosis

- changing the mode of operation and again consulting the Expert System.

The module contains the following segments, see Fig. 11:

- Mechanical data
  - Bearings
  - Thrust
  - Elongation

- Auxiliary systems

- Thermodynamic data.

In the evaluation mode, a large amount of information is available for presentation, e.g. bar charts, plant diagrams, reference curves of the set/actual value etc. It is important to note that although many values are measured, only those relevant to operation are processed and that the vast amount of remaining data is accessible for other purposes. Based on the system condition found appropriate corrective actions, stored in a knowledge base, are indicated.

Module "Characteristic Data of the Generator"

Modern generators with a high rating have a large number of measuring points (cooling water flow, voltage, current, pressures, winding temperature, etc.) which are normally used for the conventional protection of the generator (alarm and trip). Using an approach similar to that described for the module "Characteristic data of the turbine", the large amount of available measured data is condensed and compiled in functional groups where it is interpreted by the Expert System. The results including the diagnosis with warnings and actions are presented to the operator.

It should be emphasized that most of the data processed is acquired by the standard instrumentation installed in the plant. The following segments are presently being developed:

- Stator cooling water system

- Cooling gas circuit

- Seal oil system

- Mode of operation
  - Power chart
  - Rotor and bearing vibrations

- Shaft voltage, shaft current

- Excitation.

Fig. 12 shows the cooling water circuit with the most important measuring points. A measured value which exceeds the warning level indicates a change in the cooling circuit or in the generator. The target values are determined by quadratic functions which are based on so-called "fingerprints". These "fingerprints" were recorded during commissioning or after a change in the cooling system and describe the behaviour of a "sound" machine for different modes of operation.

The same method as described above (module "Characteristic operating data of the turbine") is used for storage, evaluation, analysis and representation of different parameters and for recognition of the system condition.

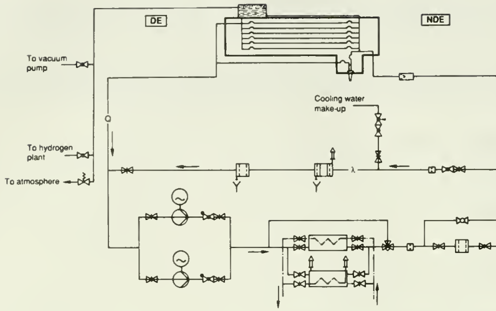


Fig. 12: Schematic view of the generator stator cooling water system (SCWS) of an ABB generator

Module "Lifetime Prediction"

In order to ensure operating reliability and high availability, on-line monitoring of lifetime consumption is recommended for all power plant components which are subjected to high pressures and temperatures and frequent temperature cycles. The determination of the actual component fatigue is of essential importance for overhaul planning and component layout. The module "Lifetime prediction" is an independent system, i.e. it does not interfere with the process and is only used for predicting the residual lifetime of HP and IP turbines. The module does not comprise an Expert System, gives no diagnosis and outputs a prognosis of the remaining lifetime.

The essential data for determining the residual lifetime of a component include details of the steam conditions, load profile, startups, shutdowns and material temperatures (3). The operating data are recorded with the lowest possible number of pressure and temperature sensors. Based on extensive studies, the conditions for the validity of a measurement and its transferability to other locations were laid down. The temperature sensors are arranged just below the steam-adjacent component surface. The radial temperature profiles in the component are calculated from the measurement signals. Fig. 13 shows the arrangement of the measuring points of a HP sliding pressure turbine.

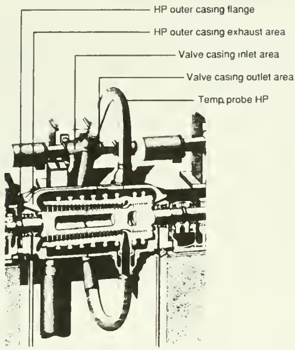


Fig. 13: Location of measuring points for determining the remaining lifetime

The lifetime consumption is determined using the criteria "creep damage" and "fatigue damage". The assessment of the "creep damage" is based on the results of the finite element methods used for the design calculations. The results are used as constants and are converted with the incoming data to the existing operating loads. The low-cycle fatigue is still determined in accordance with the Technical Rules for Steam Boilers TRD 301. The temperature cycles corresponding to the thermal stresses are calculated using software for determining the cycles according to the "rain flow range pair". The Technical Rules TRD 301, together with evaluations according to ASME and evaluations based on the results of ABB laboratory tests, are all used for calculating the stresses and the appropriate cycle temperatures. When storing the data in the long-term storage, special attention must be paid to the possibility of recalculating the remaining lifetime with updated programs.

The output comprises curves and bar charts for displaying the actual consumption of lifetime as well as a prediction of the residual lifetime. The module "Lifetime prediction" aims at higher availability and the utmost possible safety for turbine operation. The on-line system ensures fast recognition of the condition of the turbine components.

Module "Heat Rate and Performance Values"

This module is essential for attaining the main goals, as shown in Fig. 1, by optimum process control. The heat rate is a significant value of the operational state of the power plant although the parameter itself does not give the reasons for a possible deviation from the expected values. In the ABB approach the plant is divided into a number of functional groups or components, for example HP turbine, IP turbine, condenser, etc., which all contribute to a better or worse performance of the plant. This means that all components are analysed which have a marked influence on the heat rate. As in other modules, the measured data such as temperatures, pressures, differential pressures, etc. are thoroughly checked for steady-state condition. The definition of the heat rate implies a steady-state conditions in order to permit a relevant evaluation of the measured values, i.e. the data are evaluated only if the steady-state criteria are fulfilled.

With the aid of the ABB heat balance design programme, the influence factors on the heat rate are calculated and stored as functions, depending on load and cycle isolation, in the module. Using the energy balances or direct algorithms with steam tables, the performance values like turbine efficiency, condenser vacuum and heat load, LP and HP heater temperature differences and their influence on the heat rate are determined, see Fig. 14

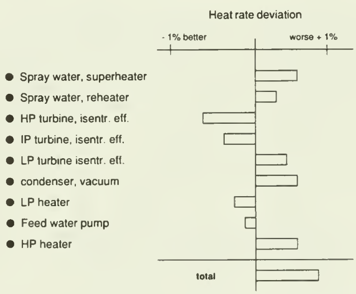


Fig. 14: Performance values and their influence on the heat rate

In addition to the heat rate, the module output informs on the condition of the components including the performance values and possible deviations from the target or reference values. The Expert System is used for interpreting performance value deviations and analyses the parameters in accordance with the preselected criteria.

CONCLUSIONS

When planning new plants the utilities are faced with the task of finding the most economic solution on a long-term basis. The owners of old plants which have been in operation for a long time must find ways to extend the lifetime of the plants. This becomes increasingly important because only a few new plants are planned and built. The On-Line Diagnostic Condition Monitoring system, based on continuous data acquisition and diagnostic evaluation, permits continuous assessment of the plant condition, contributing to the increase in the economic efficiency of the plant. One of the most important factors influencing the economic efficiency is the outage rate (forced and planned outages). On-Line Diagnostic Condition Monitoring assists the utilities in reducing the number of planned outages and avoiding unnecessary standstills of the plant. According to an estimate, the power plant availability could be raised by at least 2% by applying the most modern Monitoring technology (4).

REFERENCES

1 J.R. Walters, N.R. Nielsen
Crafting knowledge-based systems
A Wiley Interscience Publication

2 P. Harmon, D. King
Experten-Systeme in der Praxis
Oldenburg Verlag, München, Wien 1989

3 D. Heiberger, J. Schmidt, A. Czeratzki
Langzeitregistrierung von Betriebsdaten für die Lebensdauerberechnung
VDI Tagung am 17./18. September 1985 in Bochum

4 R. Scheibel
Incipient failure detection, an EPRI perspective
Presented to the third EPRI Incipient Failure Detection Conference, Philadelphia, PA/USA, March 10 to 12, 1987

EKA: An Expert System for Real-Time Operation Planning and Event Analysis in Electric Power Networks

JOUNI J. KERONEN*
Technical Research Centre of Finland
Espoo, Finland

ABSTRACT

EKA** is an expert system prototype that is intended to help operators in the control of electric power systems by facilitating switching plan configuration and checking.

EKA is implemented using object-oriented programming, rules, and temporal logic. The development environment has been the Symbolics 3645 Lisp machine, Knowledge Engineering Environment (KEE), Lisp, VAX-11/750, and Fortran.

The current prototype consists of a complete model of the 110 kV transmission network of the Helsinki Energy Board, including about 12 000 objects, 40 to 50 rules, 15 demons, a Fortran-coded power flow program, and hundreds of methods and Lisp-functions.

The first prototype was developed in Finland in cooperation with the Technical Research Centre and the Helsinki Energy Board. The work has been continued in Finland and at SRI International. A demonstration system has been installed at the Imatran Voima Ltd., the national power board of Finland.

The purpose of this paper is to describe system functions, the prototype development cycle, experience gained so far, and future plans.

* Mr. Keronen is a visiting fellow at SRI International. He will resume his association with the Technical Research Centre of Finland in August 1989.

** EKA is a Finnish acronym for an expert system for power system operations.

INTRODUCTION

With the growth of power systems, centralized control and diagnosis of power system problems are becoming increasingly difficult. Simultaneously, the rapid development of technology and increased use of electric appliances have prompted demand for enhanced quality of electricity.

The introduction of advanced information technology into power system operation has stimulated interest in more effective use of computerized analysis and control techniques. The potential uses of knowledge based systems have attracted particular attention.

Several expert systems have been developed during the pat several years years for different tasks in power system planning, control, and analysis. Because most of the systems have been based on rule-based programming [3,5,6,12,15,16] their knowledge representation capabilities have been quite narrow.

In the EKA project our goal was to explore other knowledge representation techniques and apply them to the real-time operation planning and event analysis.

REAL-TIME OPERATION PLANNING

Real-time operation planning covers numerous activities. This study concentrates on planning, generation, and testing of switching procedures. These are common activities in a power system control center, needed during all maintenance operations and recovery operations.

Switching plans are expressed in two ways: in normal situations, using switching plan forms, and in urgent situations, using a special macro command language. A simple switching plan form is represented in Table 1.

In contrast to the Table 1 example, the plans could be quite complicated. Extreme care is needed in the generation and checking of these plans to avoid the risk of incorrect ordering of switching actions which could result black out or breakage of some components, especially disconnectors. Even with the correct ordering of actions, some intermediate states in the switching process could cause overloadings and activate protection devices [7].

The major problem in switching planning is that, especially in critical situations, operators lack the time needed to thoroughly evaluate switching plans [7].

THE EKA-SYSTEM IN REAL-TIME OPERATION PLANNING

The EKA system supports operators in the generation and checking of switching plans. The process is as follows [7]:

1. The operator defines the desired final state of the power system and tells it to the EKA system using network picture, mouse, and menus. The operator can use existing high-level goals or existing lower level goals, or control the positions of switches manually.

2. The EKA system analyzes the goals and the current state of the system and generates the needed transition sequence by combining existing lower-level sequences and possible direct controls given by the operator.

3. The system simulates the transition step by step and checks intermediate states using power flow calculations.

4. The plan form and its possible negative consequences are printed out.

The primary advantage of this kind of support is that in an urgent situation the operator can concentrate on control of the situation as a whole without becoming immersed in the detailed switching sequence planning.

As a new feature we are currently developing an automatic recovery system which is based on existing switching sequences. The difference is that whereas the current system requires that the operator defines the goal state, in automatic recovery, the goal state is defined by the program itself (Figure 1). Typical tasks for automatic recovery system are recovery after total blackout or recovery of a substation.

EVENT ANALYSIS

Event analysis is needed basically for two purposes: for real-time state identification [5] and for post-mortem disturbance analysis [8,13]. The goal of the real-time state identification is to recognize the last state of the power system and predict forthcoming situations. The goal of post-mortem disturbance analysis is a careful reconstruction that helps to identify faulty components or wrong control strategies. An example of a post-mortem analysis is presented in Table 2.

Both activities involve many common characteristics, such as collection of information from multiple sources, filtering and reordering of information, and recognition and abstraction of events. The significant difference between the two activities is that the real-time state identification must occur much

Table 1

SWITCHING PLAN FORM [7]

HELSINKI ENERGY BOARD					SWITCHING PLAN (24.12.87)			
SUB-ST	kV	ORDER	ELEMENT	BUS	CTRL-DEV	OPERATION	TIME	OPERATOR
SU	110	SI	T8	A-B				
SU	110		CB1(2)		CB	1		
	110		T8	B	DC	1		
	110		T8	A	DC	0		
	110		CB1(2)		CB	0		

Table 2

A SIMPLIFIED DISTURBANCE REPORT [8]

DISTURBANCE KK 5/85 SAT 1985-08-10

Fault type	A ground fault in phases S and T developed from the ground fault of S-phase in 110 kV busbars of substation Su
Reason	A leakage of the substation roof.
Disturbance	Total blackout except the distribution areas of substations Ta and My.
Reason	The reduction in voltage insulation capability of insulators caused by moisture.
Previous state	Two lines in maintenance: Kn-Pm, Kn-Tm Energy production before disturbance 8-9 pm : Ha1 43 MWh Ha2 59
Disturbance state	Energy production during disturbance 10-11 pm : Ha1 33 MWh Ha2 0

Table 2

A SIMPLIFIED DISTURBANCE REPORT [8] (continued)

Main events

9.46 pm	Lines: Tm-Vm, Ta-Vm, My-Hn and Su-Ps disconnected. Busbar circuit breakers: Hn, Pm and Vm opened. Transformers: SuM5 and SuM8 disconnected. Generators Ha1, Ha2 and Ha4 disconnected. Blackout over the entire network except the delivery areas of the substations Ta and My.
9.50 pm	Third and fourth step distribution restriction in the 10 kV and 20 kV networks.
9.52 - 10.04 pm	Line circuit breakers: Vm-Tm, Vm-Ta and My-Hn closed. Generator Ha1 synchronized to network. Busbar circuit breakers: Hn, Pm, and Vm closed.
	etc.

Comments Far away from the Helsinki network a ground fault was noticed in R-phase. It increased phase voltages S and T and after 50 ms caused a ground fault in the busbars of Suvilahti substation. A busbar protection device indicated operation. The triggering circuits of the protection device were cut after previous operation and it did not open circuit breakers.

etc.

Suggestions If busbars of 110 kV substation should be taken into use after operations of protection devices without a complete inspection, the busbars should be used divided by groups.

FAULTS AND CIRCUIT BREAKER OPERATIONS 1985-08-10 9.46.40...45

CB OPERATIONS 110 kV	TIME/s	FAULTS
	0.00	R-phase ground fault in external network
	0.05	S-phase ground fault in Su
	(0.01)	R-phase ground fault in external network isolated
Tm Vm CB O Ta Vm CB O My Hn CB O	0.48	ground fault current 3 kA 0.5 kA
Vm Tm CB O	0.60	
etc.		

Table 3

13 TEMPORAL RELATIONS [1,2]

before x y (x, y)	after x y
meets x y	met-by x y
overlaps x y	overlapped-by x y
during x y	contains x y
starts x y	started-by x y
finishes x y	finished-by x y
equal x y	

faster, between 30 seconds and 5 minutes, while the post mortem analysis could last several days.

The major problem for both activities is that they involve manipulation and analysis of information from several sources and which is incomplete, inaccurate, and overlapping [5].

THE EKA-SYSTEM IN EVENT ANALYSIS

The aim of the EKA system is to help the operators and post-mortem analyzers to filter and organize the event information and to represent it with appropriate abstractions.

The basic idea of the system is that it has knowledge of the most typical event occurrences and their relationships as represented by procedures, processes, and event chains and that it tries to explain real-world measurement data by using these higher abstraction entities [5]. An example is given in Figure 2.

THE STRUCTURE OF THE EKA-SYSTEM

EKA is a model-based system in which the power network components and other needed structural entities are described using object-oriented programming. The behavior is described using methods, and the analytical knowledge is described using both methods and rules. The basic structure is represented in Figure 3.

SWITCHING SEQUENCE GENERATION AND CHECKING KNOWLEDGE AND REASONING PROCESS

The knowledge for switching sequence generation is represented (Figure 4) with methods divided into several layers of abstraction hierarchy [7]. The lowest level is the component level where each switch has a method OPEN! or CLOSE! whose activation will result the respective action.

At the next, or cell level, several switches are grouped to control the connections of the end of a line, a transformer, a generator, etc. Here the switching knowledge is represented with common methods, which are implemented into the subclass level in the cell hierarchy and instantiated when they are called from a cell instance. The tasks of the cell-level methods are to analyze the current switching state of a cell and organize the component level openings and closings so that the desired effect is achieved. Typical operations are changing a busbar of a transformer or a line.

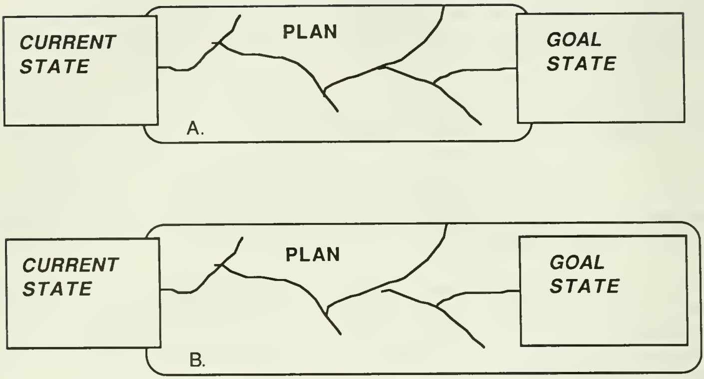


Figure 1. A comparison of a current EKA system (A) and an automatic recovery system (B)

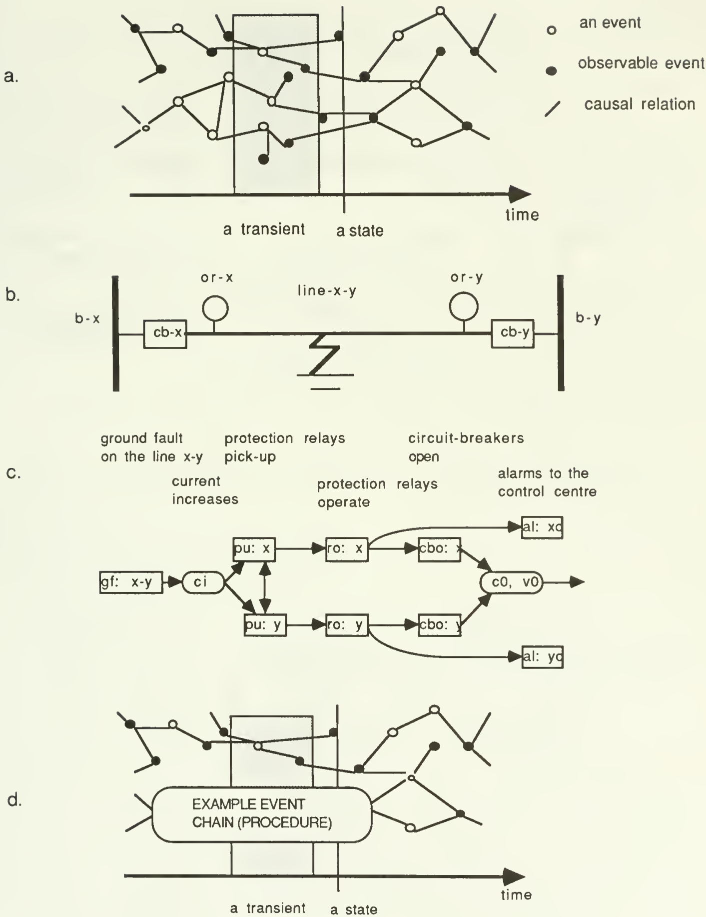


Figure 2. Pattern matching in event recognition [Keronen 1989]. A. Event data base. B. An example line configuration. C. Overcurrent protection sequence. D. Event data base after pattern matching. [8].

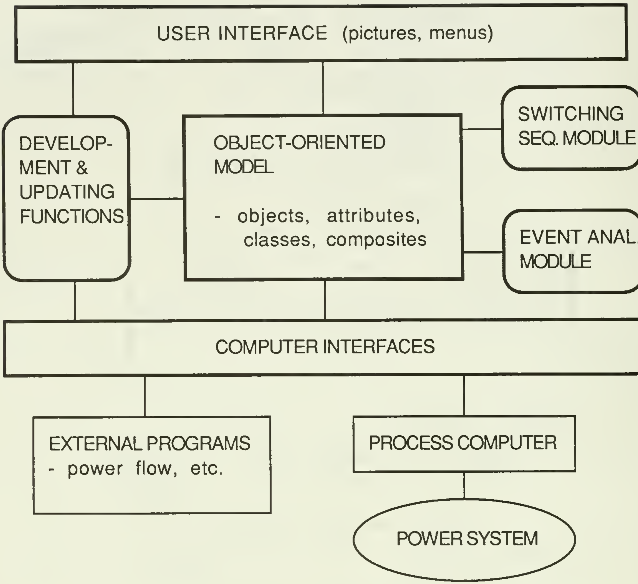


Figure 3. The structure of the EKA-system.

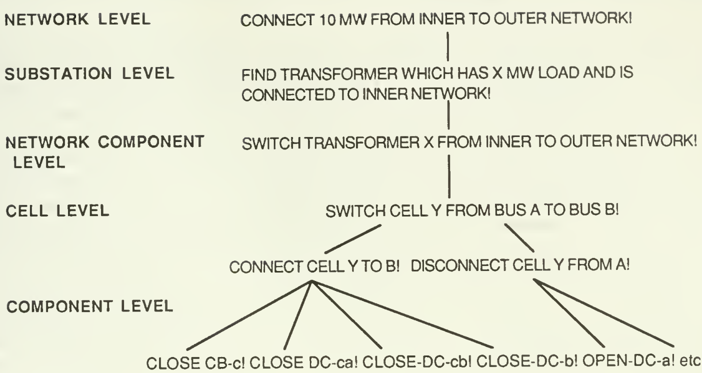


Figure 4. The hierarchies of switching methods.

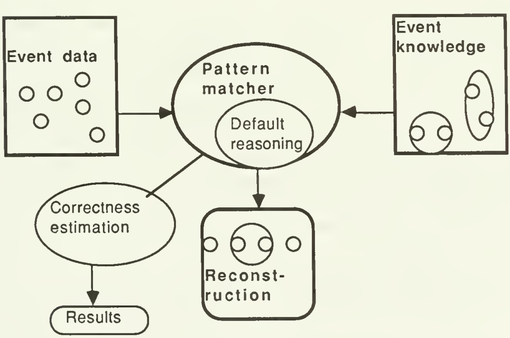


Figure 5. The event analysis reasoning process [8].

Above the cell level are a network component level, a substation level, and a network level, each with its own switching methods using lower-level methods as previously described.

During sequence generation, sequence ordering is checked with demons. This is especially important when the sequence is a combination of manual controls and existing sequences. Operations that would connect nodes with excessive voltage differences are also checked on the fly with demons.

When the plan is generated its effects on the power flow are checked by calculations after every change in the electrical state of the network. This network state (nodes, branches, isolated networks) is represented with a tree of lists which is generated and maintained with Lisp-functions. When these functions notice changes in the node structure, they send a message to the power flow calculation functions. These functions in turn create an input file and send it to the calculation computer, where the power flow program calculates the power flow and sends the results to the Lisp-computer. The results are converted into lists and analyzed by demons. The results of this analysis are printed into the switching plan and, if desired, illustrated graphically.

EVENT ANALYSIS KNOWLEDGE AND REASONING

The current version of EKA lacks event analysis knowledge. This is now under now in construction and testing phases. The primary aim is to represent the knowledge using time knowledge entities, which are:

- <u>Instantaneous entities</u>: a state, an action, a chain of states,
- <u>Time interval entities</u>: a state, an action
- <u>Mixed entities</u>: a process, a procedure.

The entities use causal, eventual, and temporal relations as their internal and external links. Causal relations are used to express why something happened or what is needed to cause something to happen [10]. Eventual relations are used to express events which would eventually occur. Temporal relations express the relationships between events in time. Currently 13 relations (represented in [1,2]) are used. See Table 3. Combinations of causal, or eventual, and temporal relations are also possible.

The reasoning has two phases: pattern matching and simulation, as shown in Figure 5. In pattern matching the existing knowledge entities are matched to the existing event data base and a new reconstructed event data base is created. In simulation the reconstructed data base is executed in a manner similar to Georgeff's Procedural Reasoning System, PRS [4]. The reasoning also includes other types of inference, such as pattern matching correctness estimation, which is planned to be done using evidential reasoning [11] and the estimation of time-incorrect process data where default reasoning [9] is going to be applied.

USER INTERFACE

The EKA system combines graphic user interface with dynamic menus and a mouse. All pictures are represented with object hierarchies similar to the components or composites. Figure 6 illustrates an end user interface.

Specialized features are the representation of critical parameters [14], Figure 6, and the planned representation of events, Figure 7.

PROJECT HISTORY AND FUTURE PLANS

The project was undertaken preliminarily in 1985 when different expert system candidates were studied and two demonstrators were implemented. In the evaluation of candidates the event analysis was seen as the most important application and the switching planning support was as second in importance. The lack of time-dependent reasoning tools forced us to start with the switching planning application; this also proved to be the easier starting point.

The first prototype of the switching planning system was completed in May 1988 and introduced to the operators in a three week training course. The course revealed that the system, particularly the analysis of the electrical state of the network, was much too slow but otherwise acceptable. Development of new algorithms for the electrical state analysis was completed in December 1988 with their integration into the system. The result was that version two was much (3-100 times, depending on the problem) faster than the first version.

In June 1988 the development team split into two parts and a new subproject was established. The main switching planning project was conducted in the Technical Research Centre of Finland with the goal of implementing more complex switching tasks, such as system the recovery from total blackout. The subproject was the idealization and feasibility study for the event analysis conducted at SRI International in California. Its goal was to apply the EKA system to the network of Imatra Power Company Ltd.

The current prototype consists of a complete model of the 110 kV transmission network of the Helsinki Energy Board, including about 12 000 objects, 40 - 50 rules, 15 demons, a Fortran-coded power flow program, and hundreds of methods and Lisp-functions.

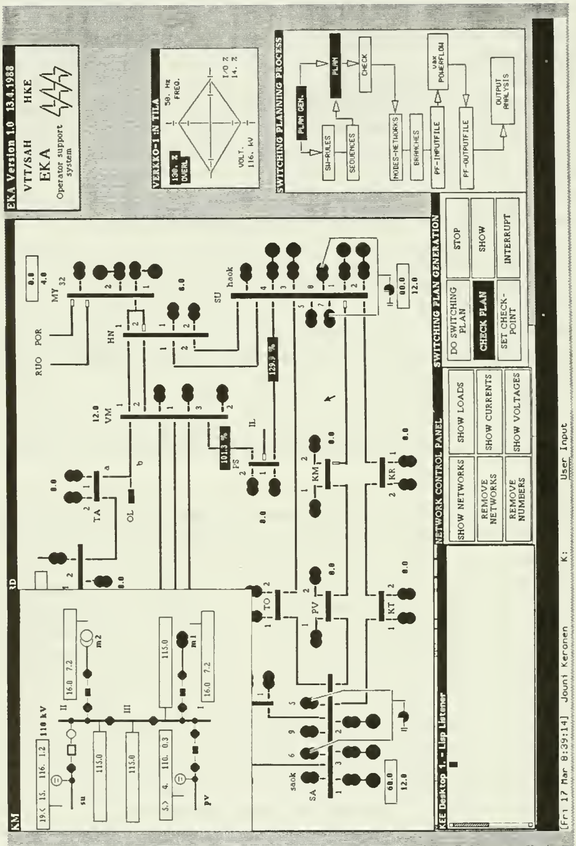


Figure 6. The end user interface of the EKA-system.

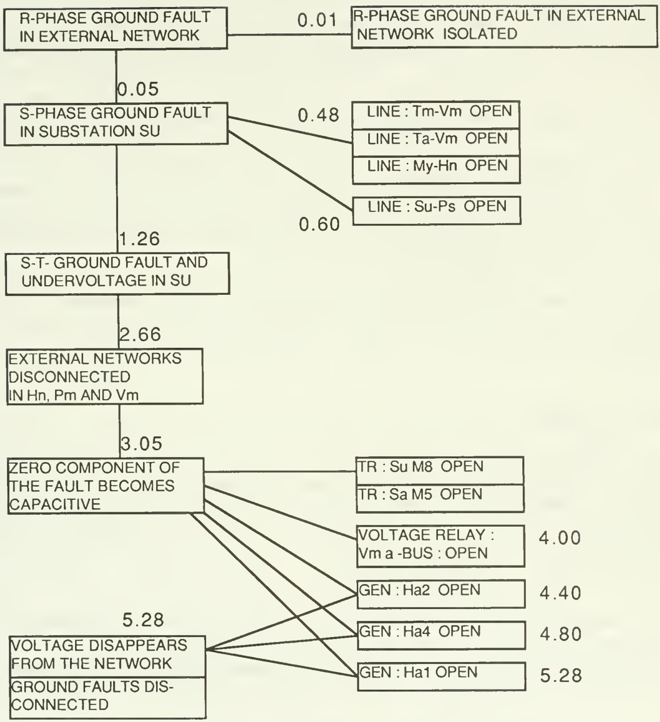


Figure 7. The planned event display [8].

The switching planning system is a waiting for testing at the Helsinki Energy Board. This should start in the next few months. The event analysis based on the idealization and feasibility studies, and some tests with a small prototype have been completed. The integration of the event analysis knowledge and the main EKA system should occur before December 1989.

The final version is intended to be installed in the control center of the Helsinki Energy Board in 1991 - 1992, when test should be complete.

The development environment has been Symbolics 3645 Lisp-computer, Knowledge Engineering Environment (KEE), Lisp, VAX-11/750 and Fortran.

So far the work has entailed some 4 man-years of effort, labor costs about $ 400, 000 and tool costs of about $ 100, 000.

The work has been financed mainly by the Finnish Ministry of Trade and Industry, supported by the Helsinki Energy Board and Imatra Power Company Ltd.

CONCLUSIONS

The model-based approach has been suitable for the problem. The object-oriented representation seems to offer a natural solution in describing power networks, and has been easy to use as a basis for analysis, diagnosis and hypothetical experiments. The flexibility and the modifiability of the user interface have made it possible to handle large numbers of entities efficiently.

The biggest problem so far has been execution speed. Use of the system in real time with response times of less than 20 seconds may not be possible with current tools. However, continuing rapid development of tools is likely to eliminate this problem within the next few years.

REFERENCES

1. Allen J. Maintaining knowledge about temporal intervals. Communications of the ACM. Vol. 26. No. 11 1983. pp. 832-843.

2. Allen J. Towards a general theory of action and time. Artificial Intelligence, Vol. 23 1984. pp. 123-154.

3. Fujiwara R., Sakaguchi T., Kohno Y., Suzuki H. An intelligent load flow engine for power system planning. IEEE PICA Conf. Japan 1986. pp. 236 - 241.

4. Georgeff M., Lansky A. Procedural knowledge. SRI International, Technical Note 411. January 1987. 32 p.

5. Gonzales A., Osborne R., Kemper C., Lowenfeld S. On-line diagnosis of turbine-generators using artificial intelligence. IEEE Transactions on Energy Conversion EC-1, No. 2, June 1986.

6. Keronen J. Expert systems in energy systems. Technical research centre of Finland. Research notes 580. Espoo 1986. 84 pp.

7. Keronen J. A knowledge based system for real-time operation planning and event analysis in power system control. Technical research centre of Finland. Research notes 883. Espoo 1988. 108 p.

8. Keronen J. A knowledge based system for power system event analysis. Preliminary version of a PhD-thesis. Unpublished. 1989.

9. Konolige, K. On the relation between default and autoepistemic logic. Artificial Intelligence 35, 1988, pp. 343-382.

10. Lansky A. Localized event-based based reasoning for multiagent domains. SRI International, Technical Note No. 423. January 1988.

11. Lowrance J., Garvey T. Evidential reasoning: An implementation for multisensor integration. SRI International, Technical Note 307. December 1983.

12. Sakaguchi T., Tanaka H., Uenishi K., Gotoh T., Sekine Y. Prospects of expert systems in power system operation. 9th Power System Computation Conference. Cascais, Portugal, September 1987. p. 29.

13. Sarmanto R. Post mortem analysis. Unpublished notes. Helsinki Energy Board. 1983.

14. Talukdar S., Cardozo E., Leao L. TOAST: The power system operator's assistant. IEEE Expert Systems, July 1986.

15. Wise J., Culver M., Abi-Samra N. Display design for dispatch control centers for electric utilities: Handbook. EPRI EL-4960, Project 2475-1, Final report.

16. Wollenberg B. Feasibility study for an energy management system intelligent alarm processor. IEEE Transactions on Power Systems, Vol. PWRS-1, No. 2. 1986.

An Expert System-based Optimal Power Flow

BADRUL H. CHOWDHURY
Electrical Engineering Department
University of Wyoming
Laramie, Wyoming 82071, USA

ABSTRACT

The optimal power flow (OPF) is fast becoming an invaluable tool for both power system planners and operators. For real-time operational purposes, an on-line implementation is required which necessitates faster execution times and minimum storage allocations. These constraints elevate the nature of the OPF problem to an extremely high level of complexity such that control centers are still quite some way from using existing techniques for real-time dispatching. The research effort of numerous authors on the problem is recognized in this paper and certain problem areas are identified. An expert system (ES) is considered as an additional tool to the power system dispatcher for rendering diagnoses and expert decisions during system insecurity. Emergency measures amount to rescheduling the power flow during branch flow violations and/or controlling the voltage and reactive power during voltage limit violations. The proposed dispatch strategy includes a full-fledged Newton's OPF executed only two to four times during the hour, an expert system invoked only during system emergencies to select control strategies for countering security violations, an economic dispatch which is executed five to six times as frequently as the full OPF and an ac power flow that is used for verification purposes.

INTRODUCTION

The optimal power flow (OPF) problem plays an extremely important role in the operation of power systems, since it calculates the power outputs and the voltage magnitudes of the generators so that the cost of power generation is minimized. In addition to the economical aspect, the OPF problem should include system security to ensure that security limits of the generators and the transmission lines are not violated. OPF problems are large-scale nonlinear optimization problems that involve the determination of the optimal steady-state operation of the electric power generation-transmission system. Optimal steady-state operation is achieved by adjusting the values of certain controllable quantities to minimize the value of a chosen objective function subject to satisfying certain equality and inequality constraints.

Real-time solutions of the OPF problem implies the minimization of instantaneous cost of active power generation on an operating power system subject to preventing violations of operating constraints in the event of any planned contingencies. Such an on-line implementation requires fast execution times and minimum storage allocations. Undoubtedly, these constraints elevate the nature of the OPF problem to a high level of complexity.

A great deal of research effort has gone into the solution of the Optimal Power Flow problem since Dommel and Tinney [1] first introduced the concept of using load flow solution techniques to the solution of the OPF problem. The method consists of extending Newton's method to yield optimal flow solutions. In this method, the incremental losses are calculated from the Jacobian ordinarily used in the Newton-Raphson load flow. The authors divide the variables into unknowns (x) which consists of (V) and (θ) on (P,Q) buses, and (θ) on (P,V) buses. Denoting the fixed parameters P,Q on the (P,Q) buses, and θ on the (P,V) buses by the parameter "p", and the control parameters as voltage magnitudes on generator buses, generator real powers, and transformer tap ratios by the parameter "u", the derivation of the authors may be summarized as

$$\min_{u} f(x,u) \tag{1}$$

subject to the equality constraints of the load flow equations

$$g(x,u,p) = 0 \tag{2}$$

the Lagrangian function takes the form:

$$L(x,u,p) = f(x,u) + [\lambda]^T \cdot [g(x,u,p)] \tag{3}$$

where λ is a Lagrangian multiplier. The set of necessary conditions for a minimum are:

$$\frac{\partial L}{\partial x} = \frac{\partial f}{\partial x} + [\frac{\partial g}{\partial x}]^T \cdot \lambda = 0 \tag{4}$$

$$\frac{\partial L}{\partial u} = \frac{\partial f}{\partial u} + [\frac{\partial g}{\partial u}]^T \cdot \lambda = 0 \tag{5}$$

$$\frac{\partial L}{\partial \lambda} = [g(x,u,p)] = 0 \tag{6}$$

Equation (4) contains the transpose of the Jacobian which can be solved for λ.

$$\lambda = (-[\frac{\partial g}{\partial x}]^T)^{-1}[\frac{\partial f}{\partial x}] \tag{7}$$

Equations (4), (5) and (6) are solved by the method of steepest descent. The basic idea is to move from one feasible solution in the direction of steepest descent (negative gradient) to a new feasible solution point with a lower value for the objective function.

Later research efforts have been mainly devoted to the improvement of convergence characteristics, the reduction of computation time and computer storage requirements. Techniques used in solving OPF as reported in the literature range from improved mathematical techniques to more efficient problem formulation. Among the mathematical techniques, some of the more important ones are the following:

i) reduced Hessian-based optimization techniques [2],
ii) successive minimum cost flow technique [3,4],
iii) modern mathematical optimization methods such as quadratic programming [5,6,7] and linear programming [8-11] techniques,
iv) P-Q decomposition [12-15],
v) constraint relaxation [16,17],
vi) quasi-Newton approach [18],

vii) Newton's method [19,20],
viii) network approach [21,22,23].

The portion of the literature referred to above mostly belong to a recent period between 1977-1988. For previous studies published prior to 1977, one should refer to [24].

The OPF problem is by nature, a nonlinear optimization problem which seeks to adjust voltage levels, power output of generators, transformer tap positions, phase shifter angle positions and switchable shunt capacitor/reactor to minimize operating costs and system losses. The usefulness of such a tool is apparent for both planning and operating purposes. For planning purposes, it should be capable of solving reasonably large-scale problems accurately in reasonable time. For operations, an on-line version should be capable of solving a smaller system accurately but with greatly reduced computing time. As with any non-linear optimization technique, there are two main drawbacks associated with the proposed solutions to the OPF problem in real-time applications: convergence and dimensionality. Algorithm convergence can be a serious drawback if the program is to be running in real-time.

Such problems encountered in the solution methodology of the OPF problem generally led to the thinking that a more efficient overall solution method needs to be developed. An Expert System (ES) approach in addition to existing solutions of the OPF problem will be a wise choice for an on-line implementation. The diagnostic capabilities of the ES will make it an efficient tool in the dispatch strategy as repeated solutions to the load flow problem will be avoided each time voltage or power constraints are violated. In the next few sections, an attempt is made to explain the working mechanisms of the ES in relation to the OPF problem.

OPF PROBLEM STATEMENT AND THE NEWTON'S METHOD OF SOLUTION

The Optimal Power Flow (OPF) problem seeks to allocate generation among the individual units and to adjust the voltage magnitudes of generators, in order to minimize the cost of power generation. In general, the OPF problem may be stated in concise mathematical notation as follows [25]:

$$\text{Min} \quad f(\bar{u},\bar{x}) \tag{8}$$

$$\text{Subject to} \quad g(\bar{u},\bar{x}) = 0 \tag{9}$$

$$h(\bar{u},\bar{x}) \leq 0 \tag{10}$$

where,

$\bar{u}$: is the control vector, consisting of all quantities whose values can be adjusted. An example of a control vector consisting only of the real power outputs P_G and the voltage magnitude V_G of the NG generators in the system is:

$$\bar{u}^T = (P_{G1}, P_{G2}, \ldots, P_{GNG}, V_{G1}, V_{G2}, \ldots, V_{GNG}) \tag{11}$$

$\bar{x}$: is the state vector, consisting mainly of the voltage magnitudes and phase angles of all the N buses in the system. These are the unknown parameters.

f: is the cost function and it is the summation of the instantaneous operating costs F_i of all NG generators, i.e.,

$$f(\overline{u},\overline{x}) = \sum_{i=1}^{NG} F_i(P_{Gi}) = \sum_{i=1}^{NG} (a_i + b_iP_{Gi} + c_iP_{Gi}^2) \quad (12)$$

where a, b, c are constants.

g: these are the typical load flow equations.

h: these are the system operating limits and they include:

a) <u>Generator operating limits</u>. For each generator, the real power output P_{Gi}, the voltage magnitude V_{Gi} and the reactive power output Q_{Gi} are restricted by an upper and lower limit.

$$\overline{u}_{min} \leq \overline{u} \leq \overline{u}_{max} \quad (13)$$

$$\overline{Q}_{Gmin} \leq \overline{Q}_G(\overline{u},\overline{x}) \leq \overline{Q}_{Gmax} \quad (14)$$

b) <u>Security limits</u>. These include transmission line loadings and voltage constraints at load buses,

$$\overline{T}(\overline{u},\overline{x}) \leq \overline{T}_{max} \quad (15)$$

$$\overline{V}_{Lmin} \leq \overline{V}_L(\overline{u},\overline{x}) \leq \overline{V}_{Lmax} \quad (16)$$

where $\overline{T}$ is the vector of branch flows and $\overline{V}_L$ is the vector of voltage magnitudes at load buses.

In generalized notation, the power flow equation for the active and reactive power injections, P_i and Q_i, at node i can be written as

$$P_i = V_i^2(g_{ii} + \sum_j t_{ij}^2 g_{ij}) + V_i \sum_j V_j T_{ij}|Y_{ij}|\cos(\theta_i - \theta_j - \phi_{ij} - \gamma_{ij}) \quad (17)$$

$$Q_i = -V_i^2(b_{ii} + \sum_j t_{ij}^2 b_{ij} + V_i \sum_j V_j t_{ij}|Y_{ij}|\sin(\theta_i - \theta_j - \phi_{ij} - \gamma_{ij}) \quad (18)$$

where,

y_{ij} = $g_{ij} + jb_{ij}$ = branch physical admittances
t_{ij} = transformer tap ratios
ϕ_{ij} = phase shift angles
V_i = voltage at node i
θ_i = angle at node i

$$Y_{ij} = G_{ij} + jB_{ij} = \text{transfer admittance of branch ij} = -y_{ij} \quad (19)$$

$$|Y_{ij}| = (G_{ij}^2 + B_{ij}^2)^{1/2} \quad (20)$$

$$\gamma_{ij} = \tan^{-1} B_{ij}/G_{ij} \quad (21)$$

The power flow mismatch equations ΔP_i and ΔQ_i for active and reactive power injections are

$$\Delta P_i = P_i - p_i \tag{22}$$

$$\Delta Q_i = Q_i - q_i \tag{23}$$

where

P_i = actual active power injection
p_i = scheduled active power injection
Q_i = actual reactive power injection
q_i = scheduled reactive power injection

SOLUTION METHOD: NEWTON'S OPF [19]

The Lagrangian for the OPF problem is formed and written in generalized form as [19]:

$$L(\overline{x},\overline{y}) = F(\overline{x}) - \sum_{i=1}^{N} \lambda_{pi}\Delta P_i - \sum_{i=1}^{N} \lambda_{qi}\Delta Q_i \tag{24}$$

where,

F = the objective function
λ_{pi} = the Lagrange multiplier for ΔP_i
λ_{qi} = the Lagrange multiplier for ΔQ_i
N = total number of buses

The problem is to find the optimal values $\overline{x}^*$ and $\overline{\lambda}^*$ such that L is a minimum. A matrix equation set is determined by using the gradient of the Lagrangian. The matrix is of the form,

$$\overline{\overline{W}}\, \Delta \overline{Z} = -\overline{g} \tag{25}$$

Elements of $\overline{\overline{W}}$ are the Hessian and the Jacobian matrices; $\Delta\overline{Z}$ is a vector of Newton corrections and g is the gradient vector.

The authors of reference 19 use an iterative technique to find the solution. The major portion of the computational effort lies in factorization and repeat solutions of $\overline{\overline{W}}$. Inequality constraints, such as the limits on dispatchable power sources, limits on variables and limits on special functions are enforced using quadratic penalty functions. The binding inequality set is then found by using special algorithms.

A new Expert System (ES) approach is introduced in this paper to overcome the "curse of dimensionality" so that an on-line implementation becomes feasible. The ES is proposed for inclusion in parallel with the solution methodology just described so that security concerns such as branch flow and voltage violations can be handled in real time. The nature of operation of such an ES is discussed next.

AN EXPERT SYSTEM AS AN AID TO THE OPERATOR

An expert system is a computer program which is capable of mimicking the problem solving behavior of a human expert from both an internal and an external point of view. The program should be capable of explaining its natural reasoning and should be able to add new information to its collection of knowledge, called the knowledge base. In narrow problem domains, expert systems can provide higher performance, equalling or even exceeding that of human experts. Expert systems have been in existence for about twenty years and are being studied within the general area of Artificial Intelligence.

At present, there are more than fifty expert systems reported to be in use and their number is rapidly increasing. Some of the original systems are widely known as DENDRAL, MYCIN, PROSPECTOR, and R1.

An expert system acts as a repository for the knowledge and skill of an expert within a particular field of expertise called the "domain". The most commonly used knowledge representation scheme is production rules. These are rules like:

IF A THEN B .

The collection of rules form the knowledge base. The knowledge base requires programs which can retrieve and manipulate the knowledge which it contains. There are three main classes of programs which operate upon the knowledge base. They are the inference engine, the explainer and knowledge elicitation tools. The inference engine uses the knowledge base and data for a particular case to infer a conclusion, in the form of a diagnosis of a fault. The program requests case data which the user can provide, and uses this with the rules, to produce a conclusion. A fundamental property of expert systems is their ability to justify and explain their reasoning. The user will need to call in the "Explainer" programs, incorporated in the inference engine. The explainer works by providing a trace of the inference engine's reasoning. The process of obtaining an expert's knowledge and presenting it in a form which is computer compatible is known as knowledge elicitation. This process is included in the category of "knowledge engineering". Figure 1 is a block representation of the parts of an expert system.

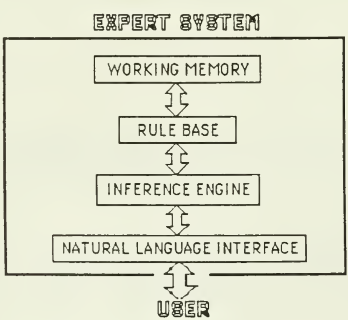


Figure 1. Parts of an Expert System.

Recently, considerable interest has been shown in the use of Expert Systems in various aspects of power system analysis, particularly in the area of Energy Management Systems (EMS) [26-30]. Modern power systems are operated by skilled operators along with the support of the EMS. Several expert systems have been developed in areas such as: load flow for system planning [31], post-fault restoration of distribution systems [32], contingency screening [33], security assessment [34] and voltage and reactive power control [35,36].

The proposed expert system is meant to be used as an assistant to the operator during times that the power system reaches a state of reduced security, or a state of emergency. In a significantly large electric utility, this situation may arise frequently. Several states of power system security have been defined by DyLiacco [37]. Transitions between one security level to a lower level is normally achieved by branch flow limit or bus-voltage limit violations. Under these circumstances where the time for action becomes of prime importance, the conventional OPF program is unable to yield proper corrective measures. The latter actions amount to rescheduling the power flow during branch flow violations and/or controlling the voltage and reactive power during voltage limit violations. An on-line implementation of the OPF program requires an additional algorithm for the corrective actions needed to restore system security. While there have been some effort in the past in generation rescheduling [38-40], no reference other than [41] is available on combining the full OPF with real-time controls. The proposed method in this paper shows how an expert system may be used in combination with a full-fledged Newton's OPF to provide real-time security dispatch.

The proposed dispatch strategy is outlined in the following steps:

Step 1: Run a Newton's OPF in a manner similar to that described in [19] by Sun, et al. The execution intervals should be between 15 and 30 minutes. This procedure should identify the binding constraints if any, as well as the set of optimal generations. The objective function to be minimized is the total cost of generation. ES is invoked if binding constraints are identified. Otherwise go to step 5.

Step 2a: Calculate the sensitivity $\overline{S}_p$ of the critical branch flow or branch current with respect to a generation change at any bus so that proper rescheduling of power may be accomplished.

Step 2b: For buses where voltage limits have been violated, determine the sensitivity $\overline{S}_v$ of the bus voltage with respect to the control measures such as transformer tap changers, switched shunt capacitors, reactors and synchronous condensers.

A simple technique introduced in [42] can be used to find the sensitivities $\overline{S}_p$ and $\overline{S}_v$. This is illustrated in Appendix I.

Step 3: The expert system determines the best possible control measure using its knowledge base and inference capability. The control actions are then taken according to certain rules, until all constraints are satisfied. In the event that certain violations cannot be overcome after using all control measures, load shedding is initiated by the ES. The operator can then decide to run a full OPF for the new operating conditions.

Step 4: After successful control measures by the ES, an ac power flow program may be executed to determine flows in all branches of the system.

Step 5: A classical economic dispatch is also executed at five to six times the frequency as the full OPF in order to determine generation levels for changes in load conditions between successive OPF runs. For the updated system configuration, sensitivity matrices are recalculated for the ES to determine any new branch flow or voltage violations. The knowledge base is updated accordingly.

Figure 2 shows a schematic diagram of the operation of the proposed expert system based optimal power flow. Flow of information between functional blocks are represented by arrows.

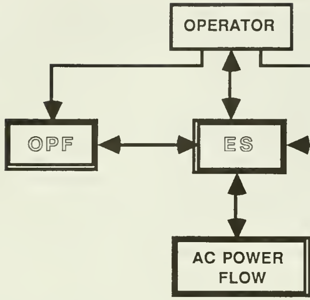


Figure 2. A Schematic of the Real-Time Implementation of the Optimal Power Flow.

BUILDING THE EXPERT SYSTEM

As described in the preceding section, the proposed expert system consists of a global data base called the working memory, a collection of rules forming the knowledge base, an inference engine and an interface for the operator to input commands or update the knowledge base.

THE DATA BASE

The data base will consist of the controlling quantities, the equality constraints and the inequality constraints. The following is a partial list:

a. active and reactive power generations
b. phase shift angles of line phase-shifters
c. transformer tap ratios
d. generator bus voltages
e. synchronous condenser outputs
f. shunt capacitances
g. bus voltage magnitudes and angles
h. branch real and reactive power flows
i. upper and lower limits of generator outputs
j. upper limits of branch flows
k. upper and lower limits of bus voltages
l. upper and lower limits of transformer tap ratios
m. upper and lower limits of phase shifter angles
n. upper and lower limits of the reactive compensators
o. sensitivity matrices or tables for each branch flow and generations at each node
p. sensitivity matrices or tables for each bus voltage and each control measure.

Note: A range of possible system operating conditions have to be considered for sensitivity matrices.

THE KNOWLEDGE BASE

The knowledge used by system operators in solving a problem consists of facts derived from physical laws and heuristics. Experience also plays a key role in strategies applied to correct the problem. For an OPF problem, constraint violations of interest are branch flows and bus voltages.

The rule base models the logic for identifying the nature of the problem and then selecting the appropriate measure for remedy. Since, the ES rule base will have many rules, a means of relating different groups of rules is required. These groups will be called "rule strands" consisting of a number of rules. All rules drawing conclusion about the state or level of system security will belong to the rule strand SA as shown in Figure 3. Branch flow and voltage are the two attributes whole values are checked for assessing system security. A modification of the security classifications of reference [37] are followed in the analysis. A normal state and three classes of the emergency state are used.

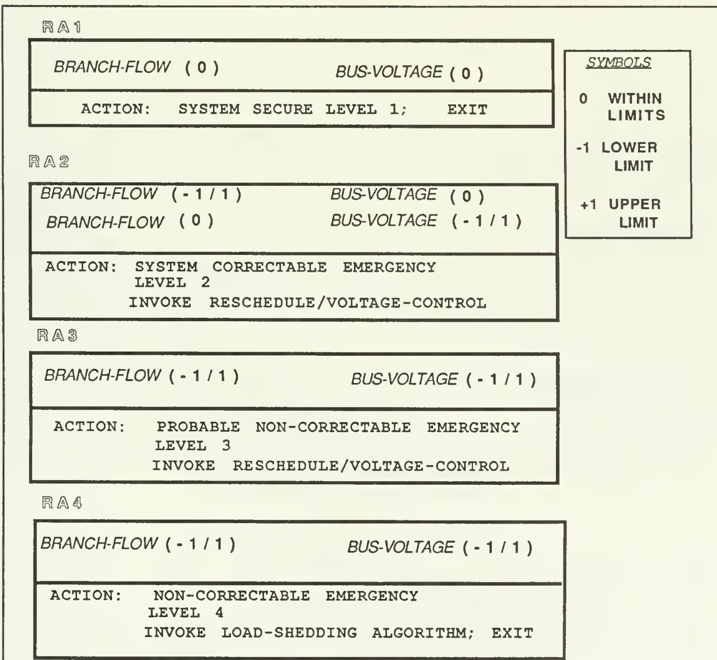


Figure 3. Partial Representation of Rule-Strand "SA".

The production rule RA2 simply states:

"If a branch is detected to be overloaded or if a load bus voltage drops below or rises above the operating limit, then the system is in security level 2."

Rule RA3 states:

"If both branch flow and voltage violations occur but affect only a number of branches or buses, then the system has attained a 'probably correctable emergency' status of security level 3; so invoke the RESCHEDULE and VOLTAGE/CONTROL rule strands."

Rule RA4 handles the case when the limit violations are too widespread over the system. The system is said to have reached a state of "non-correctable emergency".

Another rule strand called RESCHEDULE used for rescheduling real power is shown in block diagram format in Figure 4.

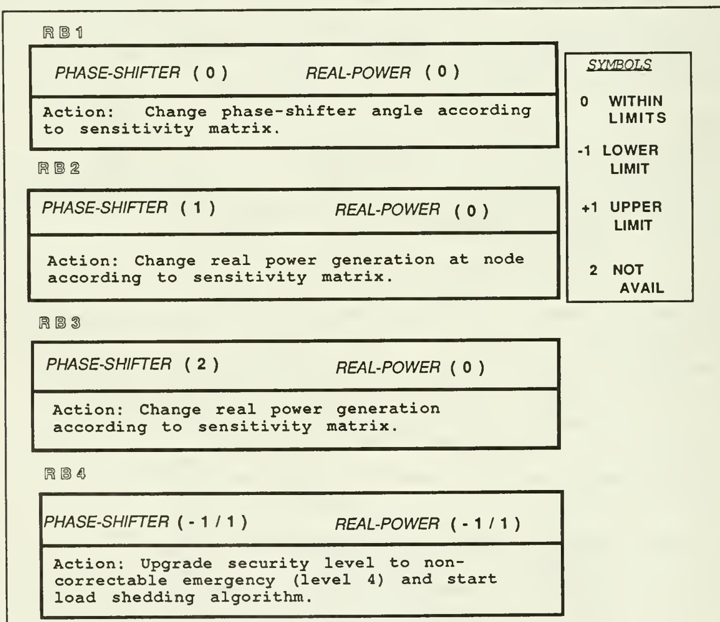


Figure 4. Partial Representation of Rule-Strand "RESCHEDULE".

Each sub-block represent rules. For example, rule RB2 would be implemented in the following manner:

> "If the phase shifter has reached its upper limit, and real power generation at nearby nodes is still within limits, then change real power generation at any node/s using the sensitivity factors of the particular branch power flow with respect to real power."

If, of course, none of the power sources nor the phase shifter in the branch are able to remedy the overloaded condition, then the security level is upgraded to level 4 of "non-correctable emergency". This is shown in the diagram at the end of rule strand RESCHEDULE. The diagram in Figure 4 is only a partial representation of the entire rule strand.

For correcting voltage problems, a rule strand called VOLTAGE-CONTROL should be developed. Figure 5 shows a possible configuration of the rules for controlling bus voltages. Once again the diagram shows a sample of rules of the actual set. Two types of controls are shown in the figure; tap changers under load (TCUL) and reactive compensators (RC). The type of controller is selected by using the sensitivity factors of the various controllers with respect to bus voltages.

CONCLUSION

The optimal power flow is characterized by exact network states and is obviously more realistic than the classical economic dispatch. The former is a proven concept in the off-line power system planning area since system planners have been using it quite successfully. However, an on-line solution of the OPF problem has consistently suffered from two main drawbacks: convergence and dimensionality. There can be serious problems if the program is executed in real time. An expert system approach is introduced in this paper to overcome the problems of on-line implementation of the OPF. The proposed ES should be used not as an alternative to the existing solution methodologies, but as an aid to the operator during decision making. The advantage lies in the fact that since the full-fledged OPF will not be running that frequently, no constraint on on-line implementation is presented. The proposed dispatch strategy includes an expert system invoked only during system emergencies, an economic dispatch which is executed five to six times as frequently as the full OPF and an ac power flow that is used for verification purposes. An aspect of security not explicitly discussed in this paper is the interaction of the optimal dispatch strategy with a contingency program so as to determine system security during contingencies. A little though reveals that the expert system can easily be used for contingency analysis as well. All that is required are some changes in the global data base to reflect changes in system condition such as line or generator outages. The ES uses these constraints and the knowledge base to either produce rescheduled generations or after exhausting all possible corrective strategies, upgrades system security to a non-correctable emergency status and invokes a load shedding algorithm.

RC1

TCUL (0)	RC (0)

Action: Check sensitivity matrices for both types of control and select 'best' type.

RC2

TCUL (0)	RC (0)

Action: If one type is not enough, select a combinationof the two by following a given strategy

RC3

TCUL (0)	RC (0)

Action: Check to see if other buses are affe-cted because of a control action by using the sensitivity factors. Use rules RC1 and RC2 to correct the problem if it exists.

RC4

TCUL (-1 / 1)	RC (0)

Action: Try the reactive compensator sequentially. If the 'best' RC has reached its limit, try another one until problem is solved.

RC5

TCUL (0)	RC (2)

Action: Try the tap-changers sequentially. If the 'best' TCUL has reached its limit, try another one.

<u>SYMBOLS</u>

0	WITHIN LIMITS
-1	LOWER LIMIT
+1	UPPER LIMIT
2	NOT AVAIL

Figure 5. Partial Representation of Rule-Strand "VOLTAGE CONTROL".

REFERENCES

1. H. W. Dommel, W. F. Tinney, "Optimal Power Flow Solutions," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-87 (10), pp. 1866-1876, October, 1968.

2. M. H. Bottero, F. D. Galiana, and A. R. Fahmideh-Vojdani, "Economic Dispatch Using the Reduced Hessian," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-101 (10), pp. 3679-3688, October, 1982.

3. T. H. Lee, D. H. Thorne, and E. F. Hill, "A Transportation Method for Economic Dispatching - Application and Comparison," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-99 (6), pp. 2373-2385, November/December, 1980.

4. T. H. Lee, D. H. Thorne, and E. F. Hill, "Modified Minimum Cost Flow Dispatching Method for Power System Application," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-100 (2), pp. 737-744, February, 1981.

5. K. Aoki and T. Satoh, "Economic Dispatch with Netowrk Security Constraints Using Parametric Quadratic Programming," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-101, (12), pp. 4548-4556, December, 1982.

6. R. C. Burchett, H. H. Happ, D. R. Vierath, "Quadratically Convergent Optimal Power Flow," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-103 (11), pp. 3267-3275, Nov., 1984.

7. M. A. El-Kady, B. D. Bell, V. F. Carvalho, R. C. Burchett, H. H. Happ, D. R. Vierath, "Assessment of Real-Time Optimal Voltage Control," IEEE Transactions on Power Systems, Vol. PWRS-1 (2), pp. 98-107, May, 1986.

8. B. Stott, E. Hobson, "Power System Security Control Calculations Using Linear Programming, Part I," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-07 (5), pp. 1713-1720, Sept./Oct., 1978.

9. B. Stott, E. Hobson, "Power System Security Control Calculations Using Linear Programming, Part II," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-97 (5), pp. 1721-1731, Sept./Oct., 1978.

10. R. Lugtu, "Security Constrained Dispatch," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-98 (1), pp. 270-274, Jan./Feb., 1979.

11. D. S. Kirschen, H. P. Van Meeteren, "MW/Voltage Control in Linear Programming Based Optimal Power Flow," IEEE Transactions on Power Systems, Vol. PWRS-3 (2), pp. 481-489, May, 1988.

12. K. Y. Lee, Y. M. Park, and J. L. Ortiz, "A Unified Approach to Optimal Real and Reactive Power Dispatch," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-104 (5), pp. 1147-1153, May, 1985.

13. X. Guoyu, F. Galiana, and S. Low, "Decoupled Economic Dispatch Using the Participation Factor Load Flow," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-104 (6), pp. 1377-1384, June, 1985.

14. R. R. Shoults and D. T. Sun, "Optimal Power Flow Based Upon P-Q Decomposition," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-101 (2), pp. 397-405, Feb., 1982.

15. G. C. Contaxis, B. C. Papadias, and C. Delkis, "Decoupled Power System Security Dispatch," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-102 (9), pp. 3049-3056.

16. M. R. Irving and M. J. H. Sterling, "Economic Dispatch of Active Power with Constraint Relaxation," IEE Proceedings, Vol. 130, Pt. C, No. 4, pp. 172-177, July, 1983.

17. B. Stott and J. L. Marinho, "Linear Programming for Power System Network Security Application," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-98 (3), pp. 837-848, May/June, 1979.

18. R. C. Burchett, H. H. Happ, "Large Scale Security Dispatching: An Expert Model," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-102 (9), pp. 2995-2999, Sept., 1983.

19. D. I. Sun, B. Ashley, B. Brewer, A. Hughes, W. F. Tinney, "Optimal Power Flow by Newton Approach," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-103 (10), pp. 2864-2880, Oct., 1984.

20. G. A. Maria, J. A. Findlay, "A Newton Optimal Power Flow Program for Ontario Hydro EMS," IEEE Transactions on Power Systems, Vol. PWRS-2 (3), pp. 579-584, Aug., 1987.

21. J. S. Luo, E. F. Hill, and T. H. Lee, "Power System Economic Dispatch via Network Approach," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-103 (6), pp. 1242-1248, June, 1984.

22. J. S. Luo, E. F. Hill, and T. H. Lee, "Bus Incremental Costs and Economic Dispatch," IEEE Transactions of Power Systems, Vol. PWRS-1 (1), pp. 161-167, Feb., 1986.

23. M. F. Carvalho, S. Soares, T. Onishi, "Optimal Active Power Dispatch by Network Approach," IEEE/PES Winter Power Meeting, Paper No. 88 WM151-3, 1988.

24. H. H. Happ, "Optimal Power Dispatch - A Comprehensive Survey," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-96 (3), pp. 841-854, May/June, 1977.

25. W. F. Tinney, D. I. Sun, "Optimal Power Flow: Research and Code Development," Electric Power Research Institute, EPRI EL-4894, Feb., 1987.

26. B. F. Wollenberg, T. Sakaguchi, "Artificial Intelligence in Power System Operations," Proceedings of the IEEE, Vol. 75, No. 12, pp. 1678-1685, Dec., 1987.

27. T. Sakaguchi, H. Tanaka, K. Uenishi, T. Gotoh, T. Sekine, "Prospects of Expert Systems in Power System Operation," Electrical Power and Energy Systems, Vol. 10, No. 2, pp. 71-81, April, 1988.

28. H. Nordgard, et al., "Knowledge Based Systems for Power System Operation," CIGRE, 1988 Session, Paper No. 39-07.

29. F. Hein, G. Schellstede, "Use of Expert Systems in Energy Control Centers," CIGRE, 1988 Session, Paper No. 39-15.

30. Y. Akimoto, H. Tanaka, "Towards Development of the Smart Systems for Power Systems Planning and Operation," Symposium on Expert Systems Application to Power Systems, Stockholm-Helsinki, Aug., 1988.

31. R. Fujiwara, T. Sakaguchi, T. Kohno, H. Suzuki, "An Intelligent Load Flow Engine for Power System Planning," IEEE Transactions on Power Systems, Vol. PWRS-1 (3), pp. 302-307, Aug., 1986.

32. C. Liu, S. Lee, S. Venkata, "An Expert System Operational Aid for Restoration and Loss Reduction of Distribution Systems," ibid, Vol. 3 (2), pp. 619-626, May, 1988.

33. D. J. Sobajic, Y. Pao, "An Artificial Intelligence System for Power System Contingency Screening," ibid, Vol. 3 (2), pp. 647-653, May, 1988.

34. R. D. Christie, S. N. Talukdar, "Expert Systems for On-line Security Assessment - A Preliminary Design," ibid, Vol. 3 (2), pp. 654-659, May, 1988.

35. C. Liu, K. Tomsovic, "An Expert System Assisting Decision-Making of Reactive Power/Voltage Control," ibid, Vol. PWRS-1 (3), pp. 195-201, Aug., 1986.

36. S. J. Cheng, O. P. Malik, G. S. Hope, "An Expert System for Voltage and Reactive Power Control of a Power System," IEEE/PES Winter Power Meeting, New York, Feb., 1988.

37. T. E. Dy Liaco, "Real-Time Computer Control of Power Systems," Proceedings of the IEEE, Vol. 62, No. 7, pp. 884-891, July, 1974.

38. K. R. C. Mamandur, G. J. Berg, "Economic Shift in Electric Power Generation with Line Flow Constraints," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-97 (5), pp. 1618-1626, Sept./Oct., 1978.

39. S. Chan, F. C. Schweppe, "A Generation Reallocation and Load Shedding Algorithm," ibid, Vol. PAS-98 (1), pp. 26-34, Jan./Feb., 1979.

40. M. Pereira, L. Pinto, A. Monticelli, "Economic Dispatch with Security-Constrained Rescheduling," Electric Power and Energy Systems, Vol. 9, No. 2, pp. 97-104, April, 1987.

41. R. Bacher, H. P. Van Meeteren, "Real-Time Optimal Power Flow in Automatic Generation Control," IEEE/PES Winter Power Meeting, Feb., 1988.

42. A. Thanikachalam, J. R. Tudor, "Optimal Rescheduling of Power for System Reliability," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-90 (5), pp. 2186-2192, Sept./Oct., 1971.

APPENDIX I: DETERMINATION OF SENSITIVITY FACTORS

The sensitivity analysis of [42] has been adopted in the determination of sensitivity factors between the controllable and the controlling variables.

The equality constraints of equation 9 is repeated here for the sake of continuity:

$$g(\bar{u}, \bar{x}) = 0 \tag{A.1}$$

where $\bar{u}$ is the control vector and $\bar{x}$ is the state vector.

Assuming that a solution $\bar{x}_0$ has been found for the set $\bar{u}_0$ Then,

$$g(\bar{u}_0, \bar{x}_0) = 0 \tag{A.2}$$

Let $\Delta\bar{x}$ be the change in the dependent variables due to a change $\Delta\bar{u}$. Hence,

$$g\left(\bar{u}_0 + \Delta\bar{u}, \bar{x}_0 + \Delta\bar{x}\right) = 0 \tag{A.3}$$

Using a Taylor's series expansion,

$$g\left(\bar{u}_0 + \Delta\bar{u}, \bar{x}_0 + \Delta\bar{x}\right) = g(\bar{u}_0, \bar{x}_0) + \bar{g}_u\Delta\bar{u} + \bar{g}_x\Delta\bar{x} = 0 \tag{A.4}$$

Using (A.2) in (A.4)

$$\bar{g}_u\Delta\bar{u} + \bar{g}_x\Delta\bar{x} = 0 \tag{A.5}$$

where,

$$\bar{g}_u = \frac{\partial g}{\partial u} \tag{A.6}$$

and

$$\bar{g}_x = \frac{\partial g}{\partial x} \tag{A.7}$$

From (A.5),

$$\Delta\bar{x} = -\bar{g}_x^{-1} \cdot \bar{g}_u \cdot \Delta\bar{u} \tag{A.8}$$

or

$$\Delta\bar{x} = \bar{S} \cdot \Delta\bar{u} \tag{A.9}$$

where

$$\overline{S} = -\overline{g}_x^{-1} \cdot \overline{g}_u \tag{A.10}$$

If the number of control variables is equal to M and the total number of dependent variables is 2N where N is the number of buses, then equation (A.9) can be written for a specific $\Delta\overline{x}$ and $\Delta\overline{u}$ as

$$\begin{bmatrix} \Delta Q_2 \\ \Delta Q_3 \\ \vdots \\ \Delta Q_N \\ \Delta V_{NG+1} \\ \vdots \\ \Delta V_N \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1M} \\ S_{21} & S_{22} & \cdots & S_{2M} \\ \vdots & & & \\ S_{2N1} & S_{2N2} & \cdots & S_{2NM} \end{bmatrix} \begin{bmatrix} \Delta P_2 \\ \Delta P_3 \\ \vdots \\ \Delta P_{M+1} \end{bmatrix} \tag{A.11}$$

The line current can be expressed as a function of the line parameters and the voltages at both ends. So, in fact a new sensitivity matrix may also be determined relating the Δ-change in line currents to the Δ-change in powers.

The branch flows are related to line currents as:

$$S_{ij} = P_{ij} + jQ_{ij} = V_i I_{ij}^* \tag{A.12}$$

where I_{ij}^* = complex conjugate of the line current.

Expert Systems for Power System Security Assessment

RICHARD D. CHRISTIE
Carnegie Mellon University and Leeds & Northrup
Schenley Park
Pittsburgh, Pennsylvania 15213, USA

PETTER STOA
The Norwegian Research Institute
of Electricity Supply (EFI)
Trondheim, Norway

SAROSH N. TALUKDAR
Engineering Design Research Center
Carnegie Mellon University, Schenley Park
Pittsburgh, Pennsylvania 15213, USA

ABSTRACT

Two problem areas limit the effectiveness of existing systems for real-time security assessment. The first is selecting the right set of contingencies to simulate. The second is interpreting the large amount of numerical information that is generated by simulating the contingencies. An off-line prototype called CQR (pronounced 'Secure') uses expert system techniques to solve these problems. It has been built and tested in conjunction with a western Pennsylvania utility. This paper describes the methods used by CQR and gives some implementation details. In particular, the use of OPS83 as the expert system shell is described.

Tests on CQR show that its reports are of comparable quality to those generated by human experts, and of far greater quality than those produced by other automatic systems. Also, CQR works fast enough to be used in real time, an order of magnitude faster than human experts can work.

In addition to its first, monolithic implementation, CQR has been implemented in a modular control framework called FORS. This framework allows easy distributed implementation and easy modification of the functional modules of CQR.

INTRODUCTION

Off-line security assessment is performed to aid in planning and maintenance scheduling, utilizing numerical tools, typically load flow programs. Engineers control the execution of these tools, provide the input data and interpret the numerical results. In on-line assessment, computer programs must substitute for the role of the engineer. Previous papers [1, 2] have pointed out how the participation of humans in off-line operational assessment produces far superior results than can be obtained by existing and fully automatic on-line techniques. These techniques can be improved by capturing the knowledge used by the humans and making it automatically available within the fifteen minute time frames typically required of real-time assessments.

One source of knowledge is the Allegheny Power System (APS), a medium sized utility in the eastern United States with interesting, non-trivial security problems. These problems stem from APS's location between midwestern coal fired generation and eastern load centers. Security at APS is affected by both internal and external events, and requires careful analysis. APS performs a daily security assessment covering the next day's operations. We have developed a computational model of this assessment process. In the model, operational security is treated as a call to action that allows for three gradations: OK (no action is needed), INSECURE (some corrective action is needed), and URGENT (immediate corrective action is needed). A tree representation models the translation from numbers describing the base case and contingencies, produced by numerical tools, to the actual security level of the power system. In the off-line assessment process, this translation is performed by the engineering supervisor of operational assessment for APS, who also selects evaluated contingencies.

Over the last two years we have been working to determine how this expert selects contingencies and how he evaluates security. This knowledge has been encoded in a rule-based program that, together with a set of numeric algorithms, constitute the hybrid expert system we call CQR. CQR has been described in [3, 4]. This paper adds discussion of the OPS83 implementation of CQR, with information on data structure and the contents of the rule base, and discussion of implementation in a framework for distributed processing.

CQR's capabilities have been growing as its knowledge base has been expanding. At present, it generates results of a quality approaching the expert's assessments (that is, far superior to the quality of a general purpose assessment algorithm), and at speeds great enough for use in real-time operations. However, actual experience with CQR in a real-time environment remains to be gained--it is still running in simulated real-time conditions.

Other expert systems dealing with security assessment are being developed [5, 6, 7], but they focus on only parts of the assessment process. CQR is believed to be the first to deal comprehensively with the complete assessment problem.

DESCRIPTION OF CQR

CQR is an expert system that uses both numerical tools and rule-based processing. CQR was originally written in OPS5, a production language developed at Carnegie Mellon [8]. CQR has been recoded in OPS83, a related production language [9], for speed and portability. This paper discusses the OPS83 version of CQR.

The numerical tools used by CQR are a fast decoupled load flow [10] and a Distribution Factors Contingency Analysis (DFAC) program [11]. These tools were originally written in FORTRAN, and recoded in C for portability in the Unix world. No significant change in performance was noted to result from the recoding.

CQR currently runs on a DEC Micro-Vax II running Unix. CQR is quite portable. It has run on Sun 3/60's running Unix, a Sun 4 running SunOS and a VaxStation 2000 running VMS. Theoretically CQR could run on any computer with compilers for OPS83 and C or FORTRAN, and a virtual operating system. CQR's memory requirements are too large for Personal Computers running MS/DOS.

CQR currently operates as an off-line prototype. Initiated from a terminal, it reads power system data from ASCII files in the PECO Power System Analysis Package (PSAP) format [12], and in some local formats. CQR then performs a security assessment using this data and writes its security reports to ASCII files.

BUILDING CQR

CQR is intended to perform an on-line security assessment task. This imposes severe constraints on tool selection. CQR's speed must be adequate for the on-line task, or a clear path for performance improvement must exist. The rule based portion of CQR must interface with numerical tools. CQR will be integrated into existing Energy Management Systems. These systems already have Human-Computer Interfaces (HCIs) that conform to specialized and stringent requirements. CQR must make use of these HCIs, not provide an additional, and different, HCI. CQR must also be portable to different hardware.

For these reasons OPS83 is used as the expert systems tool. Because it is compiled to native machine code, OPS83 has very efficient evaluation of rules, yet provides reasonable flexibility in knowledge representation and a simple yet powerful programming paradigm. It has no embedded HCI. Interfacing to functions written in C or FORTRAN is easy. It is available on a wide, but not unlimited variety of hardware, and is relatively inexpensive. The major drawback is that rule evaluation and rule syntax are not intuitive, and require some training to understand and use effectively.

OPS83 is a production system. Knowledge representation is provided in the working memory. This can contain any number of working memory elements, each containing data in a defined structure. Rules have clauses in the left hand side that form patterns. The inference engine in OPS83 efficiently searches working memory for matches to these patterns for all rules, then decides which one matched rule will be fired. When fired, the right hand side of the rule is executed, modifying working memory, and calling other OPS or external functions. This cycle repeats until no matches are found.

OPS83 has turned out to be an excellent choice. Other tools used for power system problems, at first glance far more attractive, have experienced difficulties not encountered with OPS [13].

Knowledge engineering is the process of extracting the expert's knowledge and encoding it in an expert system. For CQR, this process was performed by observing the expert at work, and asking questions about his conclusions. Initial interviews roughed out the basic structure of the system. Interviews continued at the rate of one day every two weeks until CQR could perform an assessment, although not necessarily a good assessment. Much of the time spent in this phase of development was devoted to getting the numerical tools operating properly on the APS database. Because APS uses a Newton-Raphson load flow package for operational assessment, and CQR uses a fast decoupled method, there were minor, but tolerable, problems when numerical results differed slightly due to different algorithms, and the human expert and CQR, starting from slightly different numbers, arrived at slightly different conclusions for the same power system operating state.

When CQR starting working, the visit rate was increased to one per week. During each visit CQR was run (via modem) on the same data used for the actual security assessment. The two assessments were compared and the differences discussed, in order to improve the assessment techniques in CQR. Typical time per visit was three hours, exclusive of travel.

About 150 person-days were spent over an eighteen month calendar period on CQR development, of which about 10% were spent by the expert. This time includes design and coding of the rule based program, knowledge engineering, design and coding of interfaces with the numerical tools, and resolving load flow data difficulties, but not learning OPS or coding the body of the numerical tools. The effort should be much less to implement CQR for another utility, since much of the supporting structure is now in place. However, the development should still be spread over a calendar time period of at least a year, to cover the seasonal variations in the utility's security concerns. About 25 person-days were spent translating OPS5 rules to OPS83.

A truism about expert systems is that they are never complete. Human experts continue to learn and adapt to changing conditions, and expert systems must be continually updated. Development of CQR wound down when enough success was achieved in matching assessment results to give confidence that the most important portions of the security assessment expertise at APS had been captured.

STRUCTURE OF CQR

The interface capabilities of OPS83 determine the structure of the CQR program (Figure 1). OPS83 source compiles to object modules that are compatible with the object modules produced by the C or FORTRAN compiler. OPS can call functions or subroutines contained in the C or FORTRAN object modules in the same way as it calls OPS functions, if the external functions are defined in the OPS modules. External functions, in turn, can call OPS functions and pass data to them. Both rule-based and numerical processing are contained in one program.

A small amount of utility-specific data is placed in OPS working memory when execution starts. All other data is initially read in by the numerical tools, then passed, along with numerical results, to OPS functions that create working memory elements. OPS rules create the output files.

The data structure of the OPS83 working memory is determined by the definition of <u>element types</u>. Each element type has a set of <u>fields</u>. Fields are strongly typed, that is, they must be declared to be integer, real, etc., at compile time. CQR has element types defined for each type of physical element in the power system. CQR instantiates the element type, i.e. creates a new working memory element, for each new set of data for a given physical power system element. Compared to splitting element definitions into static and dynamic components, this results in some duplication of data in working memory, but avoids combinatorial partial match problems in the inference engine. The inefficiency from data duplication has not been significant. The bus element type, for example, is:

```
type bus=element (
-- Constant portions
 number: integer;
 baseKV: real;
 hasgen: logical;  -- Set if generator attached
 genMW: real;      -- Valid only of hasgen is true
 genMVAR: real;    -- Valid only if hasgen is true
 hasload: logical; -- Set if non-zero load attached
 name: symbol;     -- Bus name
-- Variable portions
 puKV: real;       -- Computed voltage magnitude, per unit
 drop: real;       -- Computed per cent drop
 onrad: logical;   -- Bus on radial line flag
 source: symbol;   -- AC or DFAC
 caseid: integer;  -- 0 = base case
 outage: logical;  -- 1b if bus has a pre-existing outage
); -- End bus element
```

In all, there are 47 different types of elements in CQR. These may be divided into categories:

- A "goal" element type, used to control execution of OPS rules.

- Four power system data element types, "bus", "line", and two containing information about a contingency.
- Four element types related to security values. The "security_value" element type has different sub-types, one for each type of security node in the security tree.
- Sixteen element types representing intermediate results, such as counters, minimum voltage buses, MVAR sources, etc.
- Twenty two element types for constants, placed in working memory to allow access to these values from the left hand side of rules.

This data organization has proven capable of representing the data necessary for assessing security. The data representation capabilities of the OPS family of production languages have proven more than adequate for power system problems.

OPERATION OF CQR

CQR uses the procedural component of the OPS83 language to implement the major steps of the security assessment process shown in the flowchart of Figure 2. The clear boxes are implemented as C functions, and invoked by the external function call mechanism of OPS83. The shaded boxes are rule based processing, and are invoked by creating a goal in working memory to perform the function, and invoking the OPS83 inference engine.

At the start of processing, CQR invokes the AC load flow to evaluate base case operating conditions. The load flow routines read data from an ASCII file in PSAP format. This data was obtained from a seasonal planning case. Data is also read from a second ASCII file, and used to modify the power system operating state to the desired conditions. In an on-line implementation, this data would come from the Energy Management System database. Base case numerical results are transferred into working memory, and rules are invoked that evaluate base case security as OK, INSECURE or URGENT based on a tree representation of security, reflecting the view of security as a need for action, and providing some indication of the time limited nature of that need.

If there is a base case security problem, contingency evaluation is skipped, and CQR proceeds directly to writing reports. This reflects the view that there is not much value in knowing what could go wrong when something has already gone wrong. Bypassing contingency evaluation gets the security report to the operator sooner, and frees computing capacity for system response or corrective action calculations. It also imposes the requirement that CQR be absolutely correct in identifying existing security problems and suppressing false alarms.

If base case security is OK, CQR invokes the DFAC routine to evaluate real power flows for all outages internal to APS, plus selected external outages. The outage list is read from a third ASCII data file. CQR moves the DFAC results into rule based working memory, then selects AC contingencies by focusing on potential power system problems. Once AC contingencies have been selected, they are passed to the load flow routine for evaluation, and results are passed back to the expert system. The AC results replace those of equivalent DFAC contingencies. When all selected AC contingencies have been run, an explicit assessment of system security is made that includes the contingency results.

The evaluation of system security is presented on a security report. This is the way CQR communicates its conclusions to the power system operator. There are two versions of the report, operational and explanatory. The operational version is intended for real time operations. It is modeled after the written reports passed from the human security assessment expert to the operators, and is strictly limited

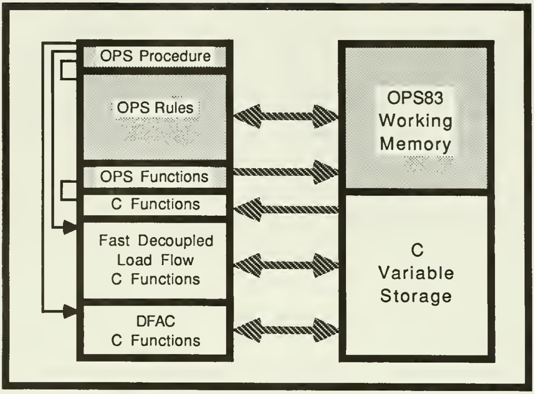


Figure 1 - CQR Software Structure

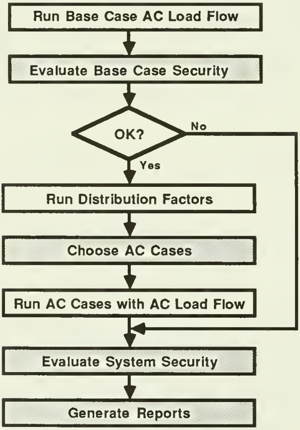


Figure 2 - CQR Operation

in length. The explanatory version is longer and contains more information. It is intended to answer questions of the form "Why did CQR think that?" when the operator has time to explore the reasoning behind the assessment.

Rule-based processing, or reasoning, in CQR is performed almost entirely by backward chaining, using goals to direct the processing of the system. There are very few forward chaining rules. This simple control structure was chosen for efficiency, and proved adequate to deal with the complexity of the problem. A goal is an element in the OPS83 working memory containing a task to be accomplished. Each type of goal that can be created has a corresponding set of rules that either accomplish the task and satisfy the goal, or create subgoals that will satisfy the original goal. Satisfied goals are removed from working memory. Initial goals created in the main, procedural component of CQR include:

- (goal type=find_case_security; value=0);
- (goal type=choose_AC_cases);
- (goal type=run_AC_cases);
- (goal type=print_reports);

For ease of maintenance, the OPS83 rule base is organized into knowledge sources. Each knowledge source contains the set of rules that deal with one type of goal. The knowledge sources have no effect on the actual operation of CQR. The rule base could be randomly rearranged without changing CQR's operation. There are 286 rules in 43 knowledge sources, giving an average of 6.6 rules each. Security evaluation accounts for 78 rules in 10 knowledge sources, 27% of the total. AC contingency selection uses 47 rules in 4 knowledge sources, 16%. Report generation uses 141 rules in 25 knowledge sources, 49%, and miscellaneous functions account for the remainder.

There are three major functions CQR provides that are not performed competently by existing assessment methods:

- Explicitly assessing security - evaluating the security tree.
- Problem focused AC contingency selection.
- Limited length result reporting.

These functions are described in subsequent sections.

THE SECURITY TREE

The concept of security is inextricably tied up with the violation of operating limits in the power system. These limits can be placed into categories. There are line loading limits, bus voltage limits, and a few additional limits on computed quantities. Separate limits apply to the base case and to contingencies. The effect of each category of limits on overall security can be considered separately. This is a decoupling, or decomposition, of the security assessment problem. This decomposition can be effectively represented in a structure termed a security tree.

CQR implements the security tree shown in Figure 3. The left half of the tree deals with the security of the the base case, and the right half with contingencies. The tree is actually a directed graph, evaluated from the bottom up. The lowest, or leaf nodes are values evaluated by numerical tools. The remaining nodes are intermediate numerical values, such as the largest EHV voltage drop, or components of power system security, evaluated as OK, INSECURE or URGENT. Each node is explicitly represented by a working memory element. The arcs of the tree are rules that evaluate the nodes, although each arc may have more than one rule.

Consider the base case (left half) of the security tree. The "Line Load Security" term is URGENT if any "Line MVA" value from the base case exceeds emergency MVA limits, INSECURE if any "Line MVA" value exceeds normal limits, and OK if no "Line MVA" value exceeds limits. Three rules - one for each possible case - are required to implement this arc in the CQR rule base.

The evaluation of voltage security at APS is somewhat complex and utility-specific. The "Voltage Security" component is derived from three intermediate values, "HV Drop", "EHV Drop", and "Hi-V Abs", the lowest absolute bus voltage on any bus with the highest base voltage in the system. Voltage drop is the difference between base case voltages and the nominal voltage profile, expressed in percent. There is one limit for EHV buses, those with base voltages over 220 KV, and a less restrictive limit for HV buses, for each of the INSECURE and URGENT bus voltage security conditions. The nominal voltage profile is recalculated seasonally, but the drop limits are constant. The Hi-V absolute limit is set independently of the seasonal voltage profile, and is usually more restrictive than the drop limits.

Buses on HV radial lines can exhibit large voltage drops. This is not considered a security problem at APS, since the problem is local and cannot develop into a system-wide condition. Even when drop limits are violated, buses on radial lines do not cause INSECURE security values. This is an example of CQR's ability to weed out false alarms that algorithmic assessment systems do not provide. Whether a line is radial depends on line switching, and must be determined dynamically for each assessment.

The set of limit violations that do not imply security problems is small. Known incorrect numerical results are the only other source. The Distribution Factors Contingency Analysis (DFAC) program, for example, can only deal with single line outages, although the arrangement of protective devices in the power system sometimes results the outage of one line causing the outage of another. Despite the small number of such situations, they occur with some frequency, and the ability to screen them out is a valuable one.

Transient stability affects operation of the APS system by imposing a limit on the sum of generator real power at one generating station. This limit is in effect only when certain lines are out of service. The limit value is determined by off-line calculations. If the limit is in force, comparison with the generation sum determines the value of transient stability security. Since violating the transient stability limit can lead to a severe system wide casualty, any violation of a transient stability limit is treated as URGENT.

Similar methods are used by other utilities to deal with the effect of transient stability on power system operations. To accommodate a wide range of similar limits, CQR provides dynamic limits. These are limits that apply to values computed from numerical values associated with one or more physical elements of the power system. They may or may not be in effect depending on power system topology, or other power system operating values. The components of dynamic limits are represented in working memory, rather than as rules. The set of operations provided to compute the limited values and the status of the limits accommodates the APS case for transient stability security, and a wide range of techniques used for applying transient stability related operating restrictions at a number of different utilities.

"Base Case Security" is evaluated by taking the worst value from its three subcomponents, "Line Load Security", "Voltage Security" and "Transient Stability Security".

The "Contingency Security" term is composed from "Contingency Case Security" terms for each contingency, that are in turn composed from "Line Load Security" and "Voltage Security" terms for each contingency. The contingencies in the security tree are those from the Distribution Factors Contingency Analysis routine (DFAC), plus selected AC contingencies. Contingency selection is not explicitly represented in the security tree. "Contingency Security" is allowed to take on only two values, INSECURE or OK, since it represents only potential, and not actual, problems. The limits for INSECURE "Contingency Security" are essentially the limits for URGENT "Base Case Security", and the voltage drop values are calculated from the base case voltages, not from the nominal voltage profile. It is therefore possible for "Contingency Security" to be OK, despite post-contingency values that, if present in the base case, would cause the system to be considered INSECURE. In operation, these situations are dealt with by corrective action after they occur, rather than by preventive action, since they present no immediate danger to the power system when they occur.

The security tree concept provides a powerful, flexible and useful way to represent and implement the explicit assessment of security. It provides a general framework for representing security, a method of discovering differences in security assessment practices among utilities, and a way to rapidly and efficiently tailor CQR to a specific utility's needs.

CONTINGENCY SELECTION

CQR selects AC contingencies by considering the types of security problems that could occur, then using heuristics to choose what is expected to be the worst contingency for each type of problem. This may be thought of as instantiating a generic problem type. Selected contingencies are evaluated with the fast decoupled AC load flow algorithm.

CQR does not use this problem focused contingency selection method for most real power problems. Complete enumeration is preferred. A Distribution Factor Contingency Analysis program (DFAC) calculates real power flows for all lines from a set of single line outages covering the entire APS internal system, plus selected external line outages. Problem focused selection could have been used to select only those contingencies that might cause real power problems, but it would take longer to pick them than it does to evaluate the complete list. DFAC can evaluate 480 single line outages in only somewhat more than the time needed for one full AC evaluation. Since the numerical tool is competent and efficient at its task, there is little justification for replacing it with rule based processing. This contrasts with the AC contingency situation, where rule based selection results in a savings in total assessment time. DFAC does not provide voltage information, and there are some contingencies where DFAC results are inaccurate. These problems are dealt with in AC selection.

APS focuses on only three problem types for AC contingency selection. The first is called transfer voltage drop. Large real power transfers through a bus can cause the voltage at the bus to drop. Increases in real power transfer cause larger drops. Large drops occurring on EHV buses are precursors to voltage collapse, and therefore of great interest to the utility. CQR looks for EHV buses where large real power transfer, while below line thermal limits, may cause excessive voltage drops. Figure 4 illustrates this situation. The EHV buses that are local minima, i.e. where all connected EHV buses have higher voltages, are located. For each such bus, the DFAC line outage causing the largest increase in real power transfer through the bus is selected as an AC contingency. Cutoffs on initial bus voltage and initial power transfer are used to limit selection to potential problems. APS views this transfer related voltage drop situation as the major security problem in their system, and it is the reason for selection of most of the AC contingencies evaluated in operational security assessment.

The second problem type is a low bus voltage caused by loss of a reactive power

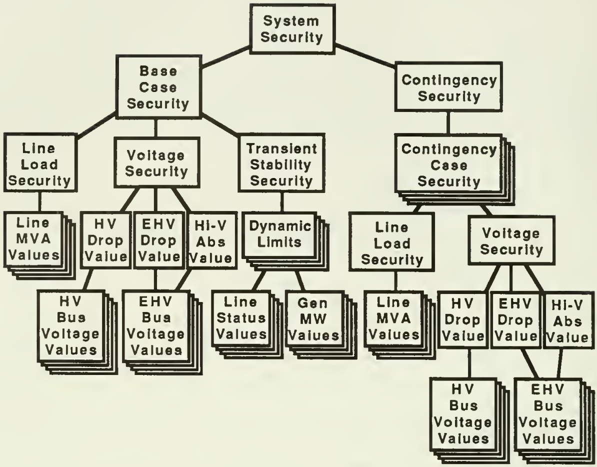


Figure 3 - Security Tree

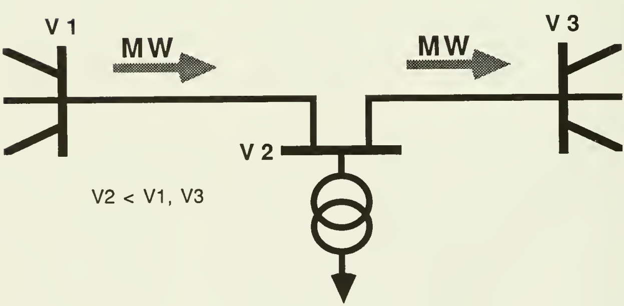


Figure 4 - Transfer Voltage Drop
AC Contingency Selection

resource (MVAR supplier). The power system is designed in the planning stage to be secure against this problem for all single outages. In addition, a good rule of thumb is that the effects of an outage, especially the voltage effects, diminish as the "distance" from the point of the outage increases. Attention is therefore focused on buses that are local voltage minima near forced or maintenance outages in the current base case. Then the largest reactive power resource supplying interesting buses is selected as an AC contingency, if voltage and MVAR value criteria for possible problems are met. Reactive power resources considered include generators as well as lines. The far segment of multi-segment lines is selected because it is a more severe problem than nearer segments. Figure 5 illustrates this contingency selection method.

The last problem type is due to inaccuracies in the DFAC results. Where there is a junction of three line segments with no circuit breakers, outage of one segment implies outage of the other two. There may also be automatic protective action that trips one line when another trips. This protective action is known as a transfer trip. The DFAC routine accepts only single line outages, so its results for these line segments may be inaccurate. This DFAC limitation is not theoretical, but rather an implementation detail. Historically, APS finds that DFAC results are accurate enough unless the line segments incorrectly remaining in service are overloaded. It is easier to run an AC contingency with all affected line segments out than to modify the DFAC program and the data representations. This situation is shown in Figure 6.

These few techniques are all those used at APS to select AC contingencies in the course of operational security evaluation. They select a small set of contingencies. Often, none of the AC contingency results have violations. The results are still of interest to the operators and used for the security report.

Problem focused contingency selection has great potential to produce security assessments with less computational effort, i.e. with fewer AC cases evaluated. The major advantage over conventional contingency screening is the elimination of evaluation of contingencies that add no new information about security, resulting in a huge savings in computational requirements. A second is the smaller set of results that still contain all the necessary information to make an assessment.

REPORTING

CQR communicates its conclusions to the power system operator via a written security report. There are two versions of the report, operational and explanatory. A key feature of the operational report is its strict length limitation. Operators can assimilate only a limited amount of information in a given time, but they always need <u>some</u> data on security. CQR respects the limit on information bandwidth while meeting the need. Existing methods do neither. This is an important and powerful feature of CQR, and a direct result of studying the human expert's methods.

Figure 7 shows the operational report for a normal operating situation, using arbitrary bus names. The report consists of three major sections, the security assessment, the base case conditions, and the contingency results. The latter section is omitted if there is a base case security problem. The assessment section is one line giving the value of security and the cause of any problem. For example, if voltage problems cause system security to be insecure, the assessment section would become:

System Security: INSECURE due to base case voltage problems.

The base case section contains a statement about transient stability, if the transient stability limit is in effect or violated, and always gives the most

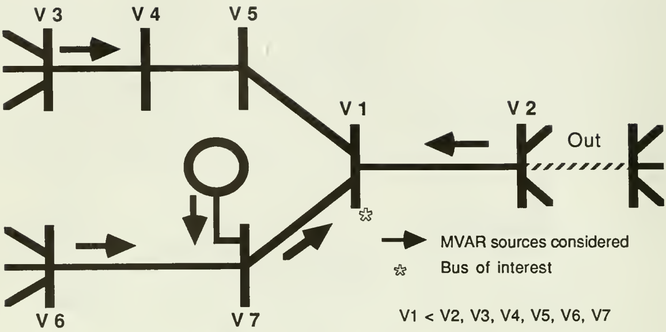


Figure 5 - Reactive Support AC
Contingency Selection

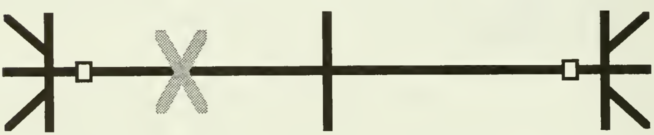


Figure 6 - Transfer Trip AC
Contingency Selection

Operational Security Report

System Security: OK

Base Case:

Bus SUBSTN A 500 voltage 512 KV (505, 500)
Line SUBSTN A 500-SUBSTN B 500 loaded to 447 MVA (550,580)

Most Critical Outages:

Loss of SUBSTN C 138-SUBSTN D 138 - 108 MVA:
SUBSTN A 500 voltage is 502 KV (500), 1.9% drop (5).
SUBSTN A 500-SUBSTN B 500 loads to 531 MVA (550,580).

Loss of SUBSTN A 500-SUBSTN B 500 - 447 MVA:
SUBSTN E 138-SUBSTN B 138 loads to 208 MVA (200,220) - over normal limit.

Figure 7 - Operational Security Report

important line loads and bus voltages. Multiple values are printed only if they are close in importance. Limits on the values are supplied in parentheses, next to the actual values. This gives the operator a feel for how close the system is to security limits, and more importantly, where in the system the problems exist or may occur. Violating values are emphasized, although only the worst violation is reported.

CQR assesses the importance of a value in different ways. Line flows use a severity index that includes the base voltage of the line, reflecting the view that security problems are more severe when they occur on higher voltage equipment. Severity is negative when the line is below limits. Bus voltages are divided into three categories. Percentage violation is compared within categories, and the categories are ordered by importance, with a violation in a category making it more important than any non-violating category. The categories are absolute 500 KV voltages, EHV (over 220 KV) voltage drop, and HV voltage drop.

Finally, the base case section may make note of operating conditions not directly related to security, such as low voltages on buses on radial lines. These voltages are reported when they are low enough to cause distribution voltage problems, and no security problems are present. They appear on the report as operating notes.

The contingency section of the operational report lists contingency results in order of importance. Each contingency is described by its outages, and lists the worst line overload, and the worst voltage, if any. Importance is a combination of heuristics and severity. The severity of a contingency is the severity of the most severe line in the contingency. Since voltage information is relatively rare, contingencies with voltages are taken as more important than contingencies without. Any contingency with a violation is taken as more severe than any contingency without a violation. However, note from the example that a post-contingency line flow exceeding normal MVA limits is not a violation. Redundant contingencies are not printed. These are contingencies with the same most severe line as some other contingency, but with less severity. The number of contingencies printed is strictly limited so the complete operational report fits on one screen of an operator display.

The corresponding explanatory report, shown in Figure 8, is an expanded and slightly reorganized version of the operational report. The report layout and the explanations allow the operator to follow the reasoning of CQR and provide a wider, but still selective, range of numerical results.

EVALUATING CQR

Some expert systems, such as those for medical diagnosis, have had elaborate and lengthy protocols established in order to attempt to objectively evaluate their quality. There has not been time to do this for CQR. Instead it is evaluated subjectively, first in comparison to operational assessment as performed by a human expert, and second in comparison to existing on-line assessment methods. The first evaluation is based on comparison with the human expert once a week over a four month period.

CQR's security assessments and reports match those of the human expert quite well. CQR identifies major security problems identified by the human expert. CQR picks about the same number of AC contingencies as the human expert, and picks the same or similar ones. CQR's reports are somewhat terser, but give the most important results with a good match to operational assessment reports. The operational report tends to have more supporting information of secondary importance, when space permits.

Explanatory Security Report

```
Base Case:

Max  HV drop at SUBSTN F 138 voltage 131 KV, 4.4% (5,10).
Max EHV drop at SUBSTN G 345 voltage 337 KV, 2.5% (3,5).
Lowest voltage at SUBSTN A 500, 512 KV (505, 500).
Absolute low voltages, EHV and HV drop are all OK.
Voltage security is OK.

Line SUBSTN A 500-SUBSTN B 500 loaded to 447 MVA (550,580)
 Severity -206.
Line SUBSTN H 345-SUBSTN G 345 loaded to 271 MVA (500,525)
 Severity -1322.
Line SUBSTN I 138-SUBSTN J 138 loaded to 201 MVA (250,275)
 Severity -1414.
No line exceeds normal MVA limits.
Loading Security is OK.

No transient stability generation limit is in effect.
Transient stability security is OK.

AC Case Selection:

Selected Case SUBSTN C 138-SUBSTN D 138:
 Possible transfer voltage problem at SUBSTN A 500.

Contingency Cases:

Loss of SUBSTN C 138-SUBSTN D 138 - 108 MVA:
 SUBSTN K 138 voltage is 132 KV, 1.6% drop (10).
 SUBSTN A 500 voltage is 502 KV (500), 1.9% drop (5).
 SUBSTN A 500 voltage is 502 KV (500).
 SUBSTN A 500-SUBSTN B 500 loads to 531 MVA (550,580).
  Severity -38.

Loss of SUBSTN A 500-SUBSTN B 500 - 447 MVA:
 SUBSTN E 138-SUBSTN B 138 loads to 208 MVA (200,220)
  Severity -369.
```

(56 more contingencies with decreasing severity values.)

```
No case is INSECURE, some case(s) are OK.
Contingency security is OK.

System Security: OK
```

Figure 8 - Explanatory Security Report

As expected, CQR is less prone to errors of omission than human beings. During testing, CQR has pointed out several mistakes made by human operators. So far, all these mistakes have been very minor. But there is always the possibility that, in the heat of the moment, an operator might forget something important which a CQR-like program would have no trouble remembering.

CQR's weaknesses in comparison to the human expert are its inability to learn from experience - it must be reprogrammed to learn - and some concern about whether enough security expertise has been captured. CQR can assess any security situation that has occurred on the APS system over the past two years as well as the human expert. The concern is over situations that have not appeared in that time, or that occur for the first time. The expertise in CQR appears fundamental enough to give confidence that very few future security problems will fall outside of its domain, although this point cannot be settled without prolonged testing.

Comparison to the human expert is important for judging how well CQR captures his expertise. The true worth of CQR, however, should be judged in comparison with existing on-line assessment methods, since this is CQR's intended domain. CQR's assessment differs fundamentally from the typical Contingency Evaluation Energy Management System software package, and is a clear qualitative improvement. This shows up best in AC contingency selection and in results presentation.

In AC contingency selection, CQR, like the human expert, picks very few contingencies. Zero to a half dozen are chosen, but these are enough to make the assessment. Current methods screen hundreds of contingencies, and perform full AC evaluation on up to fifty. CQR's advantage is that it focuses on potential problems, and picks one worst contingency for each problem, where screening methods focus on the set of most severe contingencies. This set can contain many different contingencies that cause the same problem. The CPU time spent evaluating all but the worst of these is wasted because no new information about security is obtained. CQR's selection of the worst contingency for a particular problem is an approximation. The real worst contingency may not always be picked, but the contingency that is selected will be close enough to the worst one to give adequate information about security.

The reporting aspects of CQR present more fundamental differences between it and existing on-line assessment methods. CQR makes an explicit assessment of security. Existing methods do not. CQR presents important results. Existing methods present all results, or apply a less sophisticated concept of importance, such as simple percentage overload. CQR presents important results when security is OK. Existing methods present results only when violations exist. CQR assembles the relevant information in one place. Existing methods scatter it on different displays. CQR limits the length of the results presented to the operator to an absolute maximum, by ruthlessly suppressing less important information. Existing methods do not. The estimated reduction in presented data is 10:1, improving as security degrades, since existing methods present more data to the operator as security worsens. CQR provides about the same amount of data when security is good. Existing methods often indicate good security by absence of data, giving no feel for how close the system is to problems. CQR reports in clear and understandable English language sentences. Existing methods report in tables of numbers that require an extra interpretation step to extract meaning.

Operators can assimilate only a limited amount of information in a given time, but they always need <u>some</u> data on security. CQR respects the limit on information bandwidth while meeting the need. Existing methods do neither. This concept is an important and powerful feature of CQR, and a direct result of studying the human expert's methods.

CQR's speed of execution is adequate to the real time task. The numerical tools take most of the run time, roughly 80%. Data transfer time is quite small. Performance for any combination of computer hardware and power system size can be loosely estimated by considering load flow run time. Performance is clearly adequate for on-line operation.

GENERAL APPLICABILITY OF CQR

CQR is written to perform security assessment for one utility, the Allegheny Power System. Many of the techniques used in CQR appear quite general. The best measure of generality would be to measure the effort necessary to install CQR at a new utility, and find the percentage of rules that must be changed. A faster, less expensive, but less conclusive alternative is to survey other utilities about their security practices, and estimate how well CQR could satisfy their needs. A survey of ten North American utilities was conducted on the subject of security assessment. The survey results lead to the conclusion that a surprisingly large portion of CQR is general.

The overall operation of CQR - base case, contingency selection, contingency evaluation, report generation - is common to almost all of the surveyed utilities. The exception is the use of Distribution Factors Contingency Analysis. A third used this method exclusively, a third used it in conjunction with AC evaluation, and a third used AC evaluation exclusively.

The security tree provides a general method of representing the explicit security evaluation. The tree changes in structure from utility to utility, but a tree can be drawn for each of them. Structure changes identify where new element types are needed, and where rules must be added, deleted or modified. The largest changes occur in the transition from the numerical values to the intermediate security values. The CQR method for dealing with line load security was applicable to almost all surveyed utilities. The voltage security method applied unchanged to only a third, but tree modifications to accommodate the rest were simplifications rather than complications. The transient stability security evaluation was different for every utility, but all could be dealt with, without changing rules, by redefining or adding dynamic limits.

Contingency selection is a common practice at most of the surveyed utilities. Experts "look" at the power system operating state and pick the contingencies they think might cause problems. Disappointingly, the survey did not identify any new AC selection methods, or mechanisms for problem focusing. Experts were unable to describe the techniques they used to pick contingencies in enough detail to allow replication. This inability to obtain information by direct questioning is typical of expert knowledge.

The only thing the surveyed utilities agreed on about reporting security assessment results was that very few had any formal reporting mechanism. Most often, the experts assessing security communicated verbally with the dispatchers. Dispatchers preferred short reports. Utilities disagreed on how to measure the importance of different values, when values were redundant, and what should be reported to the dispatchers.

Considering the opinions of other utilities, reporting is the least general function in CQR, and also the largest rule-based component. Yet most utilities do not have well established written reporting methods. The APS reporting techniques used to develop CQR's reporting were the only such methods found during the survey. In the absence of other established reporting methods, it is reasonable to believe the the CQR report format should be at least acceptable to a number of utilities.

In summary, CQR works well for one utility - the Allegheny Power System. It must be changed to work on another utility. CQR provides many general components that constitute a general framework for security assessment and minimize the effort required to make the necessary changes.

MODULAR CONTROL OF CQR - FORS

Re-implementing CQR in FORS (Flexible ORganizationS) is motivated by the need for a flexible, modular problem solving environment to cope with complex operational tasks.

FORS is an object oriented system intended to assemble people and programs into organizations customized for a specific task. FORS accommodates two types of objects, data objects called aspects and procedural objects called operators or tools. An aspect is a view, partial description or model of some artifact. For instance, single line circuit diagrams, transformer models and relay models are aspects of a power system. An operator is a mapping between two sets of aspects. For instance, a load flow program is an operator that maps network structure, generator settings and load values into line flows and bus voltages. FORS supports operators written in several programming languages, running in a distributed environment. It has an interface that makes it easy to execute operators and inspect aspects interactively.

CQR was split up into basic operators as shown in Figure 9. An operator is entered in FORS by stating a minimum of information about it and providing a path to its source code. The resulting graph gives a good feeling for how the assessment is performed. The graph is displayed on the computer screen and is used when interacting with the system. A pointer device is used to run operators or inspect aspects.

The FORS environment has several advantages compared to traditional EMS environments. Operators can run in parallel where possible. Every step taken when performing a task is explicit and can be examined by the users or other operators. Complex tasks can share basic operators to reduce the amount of code needed. The time it takes complex analysis programs to move from universities to utilities can be shortened by running them ad hoc until they have been proven. FORS is a promising first attempt to create an environment capable of moving complex analysis programs to the dispatcher's desk. It relies on the user to run the operators in the sequence needed to solve the problem. Automatic invocation and control of operator sequences are necessary extensions for the environment to meet on-line requirements.

CONCLUSIONS

CQR successfully addresses several major problems with on-line security assessment. The use of the security tree structure for explicit assessment of security allows inclusion of exceptions and special cases, suppressing the false alarms that result from applying the strict formal definition of security states. CQR concentrates not on the contingency set, but on the problems that the contingencies cause, and then selects the predicted worst contingency for a given problem. This drastically reduces the number of contingencies to be evaluated, and allows expansion of the reasonable contingency set to include multiple outage contingencies without greatly expanding computational requirements, since the number of contingencies selected is more a function of the number of problem types considered than the number of possible contingencies. The problem of overwhelming operators with too much numerical data placed on several different displays is addressed in CQR by the limited length security report presenting important values assembled in one location.

The problem CQR does not address is the data and software maintenance effort

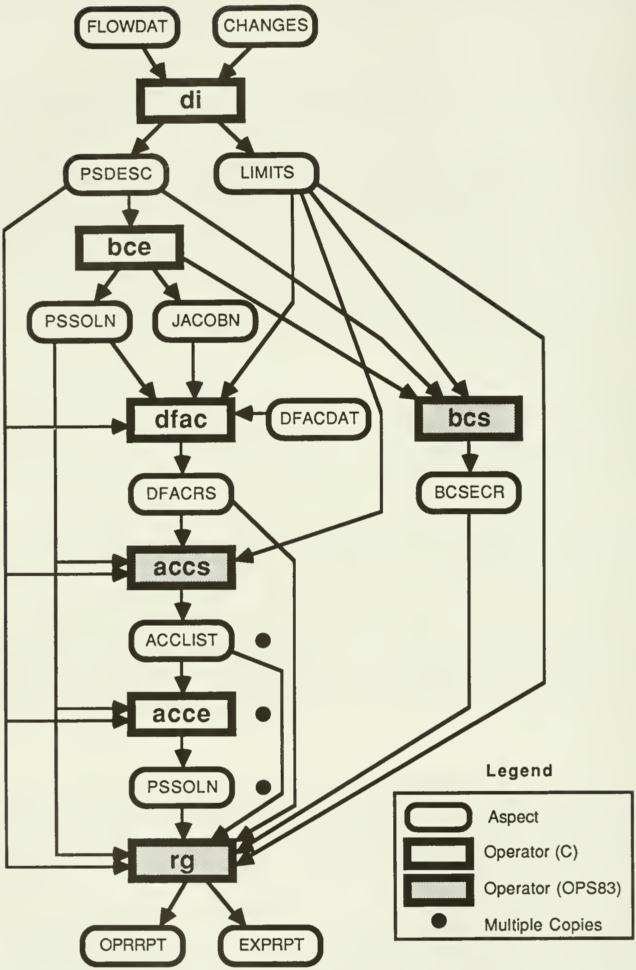


Figure 9 - Aspect/Operator Graph
for CQR in FORS

required by on-line security assessment. If anything, CQR makes this problem worse, since the data maintenance requirements of the numerical tools are unchanged, and CQR itself must be maintained. CQR at least does not require the maintenance of two separate data bases with identical information, as it gets most of its data from the numerical tools. Utility specific data in CQR is not duplicated in existing EMS databases. Maintaining CQR imposes new skill requirements on Energy Management System caretakers. It is hoped that the advantages of CQR will motivate utilities to provide adequate resources to maintain the security assessment system, and that reduction of the required effort will be a topic of future research.

CQR provides an effective means of obtaining the benefits of the security assessment expertise of human experts in the on-line environment. Its capabilities are qualitatively different from, and superior to, those of existing security assessment systems. CQR makes security assessment a useful, and more importantly, a usable function for Energy Management Systems.

ACKNOWLEDGMENTS

This work has been supported by Leeds & Northrup, a Unit of General Signal, by the National Science Foundation through the Program for Large Scale Non-Linear Systems, by the National Science Foundation through the Engineering Research Center Program, and by the Allegheny Power System.

REFERENCES

1. R. D. Christie, S. N. Talukdar, R. Edahl and J. Nixon, "Computational Sandwiches for Static Security Assessment", Proceedings of the Symposium on Expert Systems Application to Power Systems, Stockholm-Helsinki, August 22-26 1988, pp. 14.18-23.

2. R. D. Christie, S. N. Talukdar and J. C. Nixon, "CQR: A Hybrid Expert System for Security Assessment", Proceeding of the Power Industry Computer Application Conference (PICA-89), Seattle, Washington, May 1-5, 1989.

3. R. D. Christie and S. N. Talukdar, "Expert Systems for On Line Security Assessment: A Preliminary Design", IEEE Transactions on Power Systems, Vol. PWRS-3, No. 2, May 1988, pp. 654-659.

4. S. N. Talukdar and R. D. Christie, "Assessing Security Assessment", Proceedings of the International Symposium on Circuits and Systems (ISCAS-88), Helsinki, Finland, June 6-10 1988.

5. D. J. Sobajic and Y. H. Pao, An Artificial Intelligence System for Power System Contingency Screening", IEEE Transactions on Power Systems, Vol. PWRS-3, No. 2, May 1988, pp. 647-653.

6. L. Holten, A. Maeland, O. Gjerde, M. Jostad, E. Dehli and U. Wretling, "The Intelligent Operator Assistant: A Knowledge Based Security Analysis Tool", Proceedings of the Symposium on Expert Systems Application to Power Systems, Stockholm-Helsinki, August 22-26, 1988, pp. 13.1-8.

7. H. Sasaki, H. Gurita, N. Yorino, Y. Ohnogi, "An Expert System Assisting Power System Security Control", Proceedings of the Symposium on Expert Systems Application to Power Systems, Stockholm-Helsinki, August 22-26, 1988, pp. 14.1-7.

8. C. L. Forgy, "OPS5 User's Manual", Technical Report CMU-CS-81-135, Carnegie Mellon University, July, 1981.

9. C. L. Forgy, "The OPS83 Report", Technical Report CMU-CS-84-133, Carnegie Mellon University, May, 1984.

10. B. Stott and O. Alsac, "Fast Decoupled Load Flow", IEEE Transactions on Power Apparatus and Systems, Vol. PAS-93, No. 3, May/June 1974, pp. 859-869.

11. A. H. El-Abiad and G. W. Stagg, "Automatic Evaluation of Power System Performance - Effects of Line and Transformer Outages", AIEE Transactions on Power Apparatus and Systems, Vol. 81, No. 64, February 1963, pp. 712-716.

12. Pennsylvania - New Jersey - Maryland Interconnection (PJM), "PJM Power System Analysis Package User's Guide", available from Philadelphia Electric Company (PECO), Philadelphia, PA, 1986.

13. J. Keronen and M. Lehtonen, "A Knowledge-Based System for Real Time Operators Assistance in Control of Power Transmission Network", Proceedings of the Symposium on Expert Systems Application to Power Systems, Stockholm-Helsinki, August 22-26, 1988, pp. 16.1-6.

NUCLEAR POWER PLANT APPLICATIONS

CHEXPERT: An Expert System for Pipe Corrosion Evaluation

V. K. CHEXAL, J. S. HOROWITZ, and V. C. SHEVDE
Electric Power Research Institute
3412 Hillview Avenue
Palo Alto, California 94304, USA

T. C. KESSLER
Expert-EASE Systems, Inc.
1301 Shoreway Road
Belmont, California 94002, USA

ABSTRACT

Corrosion in power plants is a significant problem. Plant availability losses related to corrosion are in the range of 8-10%. In addition, corrosion raises severe plant and personnel safety concerns. In light of these issues, the challenges to EPRI were (i) to identify probable causes of corrosion, (ii) to find ways to determine where corrosion most likely has occurred in piping, (iii) to define accurate and low-cost methods to carry out inspections and (iv) to identify techniques for preventing further pipe degradation.

To address these challenges, EPRI is developing CHEXPERT, an expert system for pipe corrosion evaluation. CHEXPERT uses a combination of classical programming and expert systems techniques to provide advisory and diagnostic services related to in-service degradation of piping systems. In addition, CHEXPERT provides a training feature to educate the user in various aspects of corrosion, such as history, theory and practical solutions.

CHEXPERT considers single- and two-phase erosion, cavitation, microbial-induced corrosion (MIC) and intergranular stress corrosion cracking (IGSCC). For each of these mechanisms, the user can (i) obtain a tutorial presentation on the causes, symptoms and consequences of that mechanism along with the possible remedies, (ii) select a plant subsystem and obtain an evaluation of its susceptibility or (iii) enter appropriate information and obtain an evaluation of the probable cause of and a recommended solution for a specific problem. In addition, CHEXPERT provides a list of EPRI reports, products and contacts that can be utilized to obtain additional assistance or information.

This paper describes the capabilities, architecture, knowledge base structure and inferencing techniques used in the CHEXPERT expert system. It also provides a description of CHEXPERT's man-machine interface as illustrated by an example CHEXPERT consultation session.

INTRODUCTION

Corrosion in power plant piping systems is a complex phenomenon which depends on the interrelationship of a variety of design and process parameters including water temperature, water chemistry, piping material, fluid velocity and the geometry of the flow path. A thorough understanding of these phenomena is essential to enable power plant engineering personnel to recognize the potential for in-service piping degradation and prevent the occurrence of catastrophic piping failures. However, such broad-based knowledge spanning several engineering disciplines is rarely available among the engineering staff at a typical power plant and most likely exists only in the form of the collective knowledge of a small group of experts who have devoted extensive time to study a specific corrosion problem.

Accordingly, the Nuclear Power Division of the Electric Power Research Institute (EPRI) has formed a team of such experts and has begun the process of implementing their collective knowledge into a series of computer software products for the utility industry. The first set of products in this series, CHEC[1] and CHECMATE[2], are analytical programs which enable utility personnel to quantify the degreee of piping degradation from single-phase and two-phase erosion corrosion respectively. The codes predict wall thinning in carbon steel piping in power plants and predict the remaining service life for the piping components. These codes perform complex chemical and thermodynamic calculations for evaluating erosion-corrosion phenomena under conditions of steady single-phase and two-phase flow. Therefore, effective utilization of these codes requires a basic understanding of the physical processes which influence erosion-corrosion. However, neither code addresses the basic problem of how to make this pre-requisite knowledge available to plant personnel who don't have direct access to EPRI's team of experts. CHEXPERT is being developed to help the plant engineer to recognize, understand and identify the possible solutions for a specific corrosion problem.

CHEXPERT combines Artificial Intelligence (AI), classical analytical programming and database management technology to compile a broad base of theoretical and practical corrosion expertise. The resulting compilation is combined with EPRI's latest user interface standard (EPRIGEMS[3]) to form a Corrosion Advisor. This provides the latest corrosion technology accessible at any time to interested utility engineers. The goal of CHEXPERT is to provide sufficient insight into the physical phenomena and operational considerations that influence in-service piping degradation to enable a typical power plant engineer to:

1. Learn about various types of corrosion and how plant design and operational characteristics affect its occurrence;

2. Identify areas that are susceptible to in-service degradation;

3. Recognize and diagnose symptoms of various forms of corrosion;

4. Obtain situation-specific recommendations for preventive or corrective actions;

5. Identify and access EPRI reports, products and contacts that can be consulted for more detailed information about a particular problem.

Figure 1 identifies the various advisory services provided by the CHEXPERT Corrosion Advisor. Such an advisor would help the engineer make knowledgeable decisions for mitigating corrosion problems in the plant.

REQUIREMENTS OF A CORROSION ADVISOR

For the Corrosion Advisor to achieve these goals, it must perform certain basic tasks. These include storage and retrieval of information, obtaining and evaluating information from the user and generating meaningful reports. In addition, it must perform these tasks without intimidating or overwhelming the user with its operational complexities.

The Corrosion Advisor thus consists of:

1. A database for storage and retrieval of information;

2. A knowledge base and inference engine for evaluating information;

3. A user interface for integrating items 1 and 2 and for generating reports.

Each of these components in turn must satisfy additional requirements to function effectively, as described below.

Requirements for Database

A Corrosion Advisor database must be capable of storing and retrieving the following types of information:

1. General plant descriptive data including:
    a. The name of the unit;
    b. The type (e.g., PWR, BWR, etc.) of the unit;
    c. The subsystem of interest at that unit.
2. Metallurgical information, including:
    a. Piping material;
    b. Weld material;
    c. Cladding material, if any.
3. Hydrodynamic information, including:
    a. Primary fluid (e.g., water, steam, two-phase, oil, etc.);
    b. Fluid properties (e.g., temperature, flow rate, etc.);
    c. Flow path geometry (e.g., bends, tees, valves, etc.).
4. Operational information, including:
    a. Unit and subsystem operating history;
    b. Inspection procedures;
    c. Inspection frequency.
5. Water chemistry information, including:
    a. Treatment type (e.g., ammonia, morpholine, etc.);
    b. pH levels;
    c. Dissolved oxygen levels.

6. Descriptive information about corrosion and its effects, including:

   a. Physical processes which produce corrosion;

   b. History of corrosion in power plants;

   c. Symptoms and consequences of corrosion, supplemented by graphic displays where available;

   d. Preventive and corrective measures.

7. Lists of EPRI reports and key technical contacts for obtaining additional information on corrosion.

Requirements for Knowledge Base

The Corrosion Advisor knowledge base must be capable of processing the information described above and reasoning about it. In order to satisfy the goals of CHEXPERT, the knowledge base must be capable of:

1. Evaluating user-supplied plant data to identify whether or not a corrosion problem exists and, if so, what type of corrosion and in what location;

2. Seeking out and processing such data as is required to evaluate the susceptibility of a particular plant sub-system to various corrosion mechanisms.

In addition, the Corrosion Advisor knowledge base must be modularized to enable each of the corrosion mechanisms to be treated collectively or individually.

Requirements for User Interface

The requirements for the Corrosion Advisor user interface are that it be:

1. Visually interesting, with sufficient use of color graphics to promote active and frequent useage;

2. Self-guiding, with extensive use of menus, data entry forms and on-screen help to promote effective useage;

3. Consistent with appropriate industry "look and feel" standards to promote rapid user familiarization and acceptance;

4. Accessible on common industry computer hardware to promote widespread acceptance and useage.

CHEXPERT ARCHITECTURE

The CHEXPERT software design is governed by the EPRIGEMS software development standards. Under EPRIGEMS, a software application is constructed in a two-level hierarchy, the upper level being a generic man-machine interface (called the Session Manager) and the lower level being the specific features of the particular application. In CHEXPERT, this lower, application-specific level is further subdivided into a third level in order to support separate but parallel treatment of the five corrosion mechanisms that CHEXPERT considers.

The following sub-sections provide descriptions of the features and functions implemented at each of the three levels. The CHEXPERT architectural hierarchy is depicted graphically in Figure 2.

Session Manager Level

The Session Manager is the primary man-machine interface for all EPRIGEMS applications and defines the "look and feel" aspects of all application-specific features that lie under it. In CHEXPERT, the Session Manager level controls all user activities that are not directly related to a corrosion advisor consultation. These activities include:

1. General data and file management;

2. Tutorial about EPRIGEMS;

3. Module development and update facilities;

4. Access to external routines or other EPRIGEMS modules.

In addition, the CHEXPERT Session Manager provides mechanisms for quick access to several overview features that are specific to the Corrosion Advisor application, including:

1. Tutorial about CHEXPERT;

2. Access to the CHEXPERT reference glossary/index.

In many EPRIGEMS applications, expert system technology is utilized at the Session Manager level to guide the user through the session and to support the process of problem identification and selection of the appropriate problem solution approach. However, in CHEXPERT, this process is performed at the Corrosion Advisor level (see below) so no expert system interface is provided at the Session Manager level.

For the CHEXPERT application, the EASE+[4] graphics user interface software was used to develop the session manager and all lower levels of the application hierarchy. EASE+ was selected because:

1. It had already been used to develop the man-machine interface for CHECMATE and was therefore familiar both to the application development team and to plant personnel involved in corrosion evaluation;

2. It complies with all EPRIGEMS specifications.

3. It satisfies the database and user interface requirements identified for the Corrosion Advisor.

Corrosion Advisor Level

The Corrosion Advisor level is the second level of the CHEXPERT hierarchy and is accessed from a menu at the Session Manager level (Figure 3). The Corrosion Advisor level is the starting point for all corrosion advisor consultations and provides access only to features that are specific to the Corrosion Advisor application.

The purpose of this level is to serve as a session manager for corrosion advisor activities. The primary function of this level is to assist the user in identifying which of the five corrosion mechanisms (single phase erosion corrosion, two-phase erosion corrosion, cavitation corrosion, MIC or IGSCC) is to be investigated. When the user first

EPRIGEMS SERVICES:

DATA/FILE MANAGEMENT   DEVELOPER TOOLS   ACCESS TO CORROSION ADVISOR

CORROSION ADVISOR SERVICES:

CORROSION REFERENCES   CORROSION DIAGNOSTICS   ACCESS TO MECHANISM ADVISOR

MECHANISM ADVISOR SERVICES:

TUTORIALS   SUBSYSTEM SUSCEPTIBILITY   SITUATION EVALUATION   REFERENCES   QUANTIFICATION

Figure 1: CHEXPERT Advisory Services

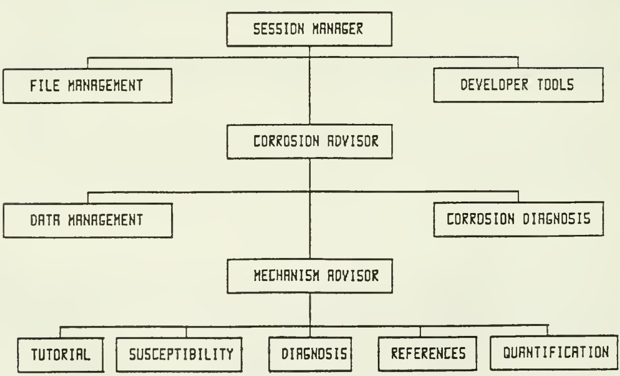


Figure 2: CHEXPERT Structural Hierarchy

Figure 3: CHEXPERT Session Manager Menu

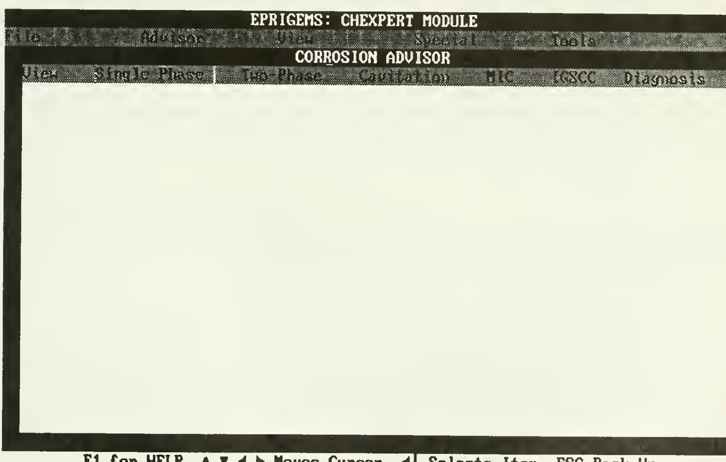


Figure 4: CHEXPERT Corrosion Advisor Menu

enters the Corrosion Advisor level, he is presented with the Corrosion Advisor menu bar as illustrated in Figure 4. The first selection in this menu provides access to the same CHEXPERT tutorial, database and glossary/index facilities that were available from the Session Manager level. The next five options allow the user to select which of the five corrosion mechanisms for investigation. This is performed by selecting the appropriate mechanism from the Corrosion Advisor menu bar, at which point control of the session is transferred to the appropriate sub-module of the next level of the CHEXPERT hierarchy for further processing.

The final selection in the Corrosion Advisor menu accesses the Corrosion Advisor diagnostic knowledge base. The purpose of this diagnostic feature is to assist the user in performing a qualitative evaluation of potential corrosion-related problems at his specific power plant. It assists in identifying which of the five corrosion mechanisms is the most likely candidate for further evaluation. After selecting this option, the user is asked to supply additional information (e.g., plant name and type, chemistry and metallurgy, operating history, etc.) that is evaluated by the knowledge base in order to select the leading corrosion mechanism. Once this mechanism has been identified, control of the session is transferred back to the Corrosion Advisor menu, from which the user can select the appropriate sub-module of the next level of the CHEXPERT hierarchy for a more detailed evauation if desired.

For the CHEXPERT application, the NEXPERT[5] expert system software was used to develop the Corrosion Advisor diagnostic knowledge base and all mechanism-specific knowledge base sub-modules used at lower levels of the application hierarchy. NEXPERT was chosen because:

1. It is the most powerful expert system software available for use on personal computers and satisfies all of the requirements for information processing listed earlier;

2. A standard information transfer protocol between NEXPERT and EASE+ had already been developed and could be applied directly to CHEXPERT, thereby reducing the overall CHEXPERT development effort.

3. It complies with all EPRIGEMS specifications;

The structure and content of the CHEXPERT Corrosion Advisor diagnostic knowledge base and all lower-level knowledge base sub-modules is described in a later section.

Mechanism Advisor Level

The Mechanism Advisor level is the lowest level of the CHEXPERT hierarchy. The purpose of this level is to provide the following specific advisory services related to each of the five corrosion mechanisms that are considered by CHEXPERT:

1. Tutorial about the selected corrosion mechanism;

2. Evaluations of the relative susceptibility of various plant sub-systems to the selected corrosion mechanism;

3. Evaluations of situation-specific corrosion problems and recommendations for corrective/preventive actions;

4. References related to the selected corrosion mechanism;

This level consists of five parallel modules, each of which provides identical corrosion advisory services for the specific corrosion mechanism selected at the Corrosion Advisor level. In addition, for flow-assisted corrosion mechanisms (single phase and two-phase) only, the CHEXPERT Corrosion Mechanism Advisor level provides access to the CHEC and CHECMATE corrosion analysis programs to allow users to perform quantitative analyses. Example results of such analyses are also provided for these two mechanisms.

Within each of the five mechanism-specific sub-modules, expert system technology is used to support one or more of the individual advisory services listed above. However, the approach taken by each module varies somewhat depending upon the nature of the mechanism and the available information about it. For example, flow-assisted corrosion is a process for which the underlying physical processes are well understood, and a wealth of quantitative information is available from CHEC and CHECMATE analyses performed under a wide variety of plant configurations and operating conditions. Accordingly, much of the information in the single- and two-phase corrosion advisor modules is quantitative in nature and expert system technology is used primarily to support quantitative analysis by relating existing data to situation-specific evaluations. However, for MIC, very little quantitative analysis has been performed and most of the available information relates to qualitative and subjective evaluation based upon system operating history and direct observation. In this module, expert system technology is used as the primary evaluation methodology for all of the advisory services.

The following subsections describe the features of each mechanism-specific advisor module and the extent to which expert systems technology is employed in support of the various advisory services provided. The Single-Phase Corrosion Advisor module is used as the primary illustrative example, and other modules are then compared to this module regarding treatment of specific features.

Single-Phase Corrosion Advisor

For flow-assisted corrosion, the physical processes involved are reasonably well understood and have been quantified using the CHEC corrosion analysis program. Therefore, most of the information presented is quantitative in nature and relates to corrosion rates that have been determined for typical power plant chemistries, geometries and operating conditions. Information contained in this module was obtained primarily from References 1 and 5.

In the Single-Phase Corrosion Advisor sub-module (and all other mechanism-specific sub-modules), the user selects the particular advisory service desired from an Advisory Service sub-menu as shown in Figure 5. The Tutorial selection provides access to detailed background information about key aspects of single-phase flow-assisted corrosion, including:

1. Underlying physical processes;

2. History of occurrence in power plants;

3. Symptoms and consequences;

4. Typical preventive/corrective measures;

This information is presented via a series of screens through which the user may page freely. In order to provide maximum flexibility, a Tutorial Services sub-menu (Figure 6) is provided to enable the user to select the full tutorial or any specific subject as desired. This service is a display-only feature with no utilization of expert system technology.

The Susceptibility selection provides an evaluation of the relative susceptibility of various plant sub-systems to single-phase flow-assisted corrosion. When this option is selected, the user is asked to select the sub-system of interest by pointing to the

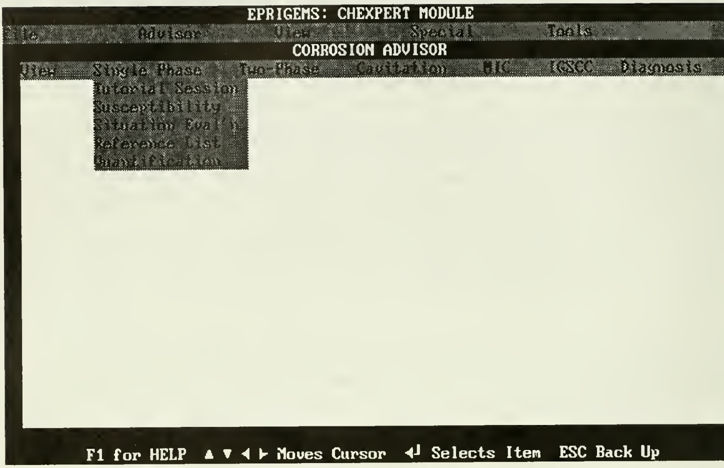


Figure 5: CHEXPERT Single-Phase Corrosion Advisor Sub-Menu

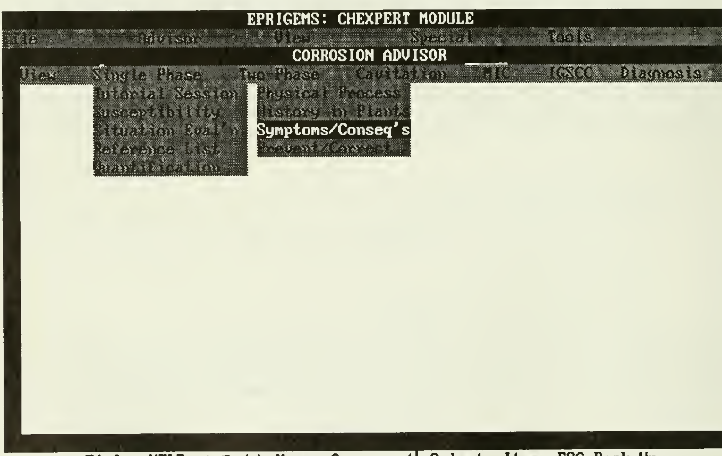


Figure 6: CHEXPERT Single-Phase Corrosion Tutorial Sub-Menu

appropriate location on a schematic diagram of a typical power plant (Figure 7). After the sub-system has been selected, the user is asked to provide more detailed information about the design and operation of that sub-system. This information is then evaluated by the Corrosion Advisor diagnostic knowledge base susceptibility sub-module to obtain a qualitative evaluation of the susceptibility of the selected sub-system to single-phase flow-assisted corrosion. This selection causes the knowledge base to be processed in a goal-driven (backward chaining) mode, while the Diagnostic option processes it in a data-driven (forward chaining) mode. The results of this evaluation are presented in the form of a qualitative susceptibility rating (e.g., High, Moderate, Low) accompanied by an explanation of the specific design and operation parameters that supported that rating. Figure 8 shows a typical susceptibility evaluation rating and explanation display.

The Situation-Specific Evaluation selection determines whether or not a single-phase corrosion problem actually exists and, if so, what should be done to correct the situation or to prevent further degradation. This module is a more detailed version of the general Corrosion Advisor diagnostic option and attempts to pinpoint the location and severity of a specific problem rather than identifying only the most likely corrosion mechanism. As with the Susceptibility selection described above, the user is asked to supply additional design and operation information which is processed by a sub-module of the Corrosion Advisor diagnostic knowledge base. However, this selection, like the general diagnostic option, processes the knowledge in a data-driven mode. The results of this evaluation are a ranked list of possible corrosion problem areas accompanied by appropriate recommendations for corrective/preventive action. Figure 9 shows a typical results display for a situation-specific evaluation.

The References selection provides access to a glossary of key terms and definitions associated with single-phase flow-assisted corrosion, together with a reference list of EPRI reports, products and contacts that can be consulted for additional information. This selection is a sub-set of the overall CHEXPERT glossary/index and reference list that is available at both the Session Manager and Corrosion Advisor levels of the CHEXPERT application hierarchy.

The Quantitative Evaluation selection, which is limited to only the single-phase and two-phase corrosion sub-modules, provides access to the results of quantitative analyses obtained from sample cases of the CHEC corrosion analysis program. When this module is selected, the user is asked to select the plant type and configuration that most closely resembles his own plant from a list of "typical" configurations that have been analyzed by CHEC. He is then presented with the results of sample calculations for representative geometries within that configuration. If the user is also a CHEC/CHECMATE licensee, this option also provides direct access to these codes to perform new analyses as required. Figure 10 illustrates typical CHEC/CHECMATE analysis output as displayed by or generated by this selection.

Two-Phase Corrosion Advisor

The Two-Phase Corrosion Advisor sub-module is identical in both form and function to the Single-Phase Corrosion Advisor described in the previous sub-section. Both modules provide the same features in the same format and use expert system technology in the same manner. The only differentiating factor is that the two-phase module addresses only those plant sub-systems in which steady two-phase flow or flashing is likely to occur, and the specific operating parameters requested by the Susceptibility and Situation-specific Evaluation selections include additional parameters relating to two-phase flow conditions.

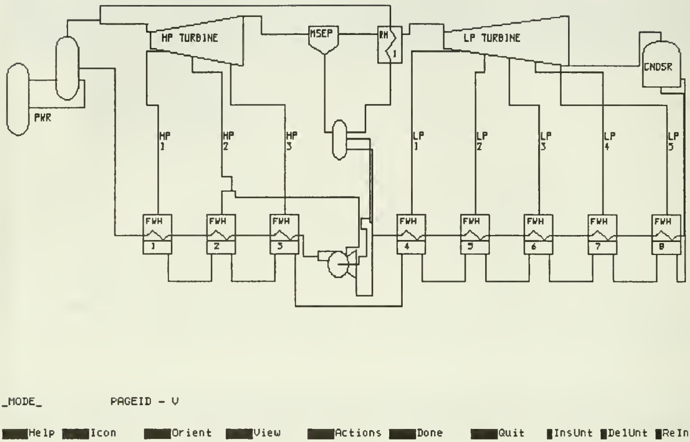


Figure 7: CHEXPERT Diagram for Sub-System Selection

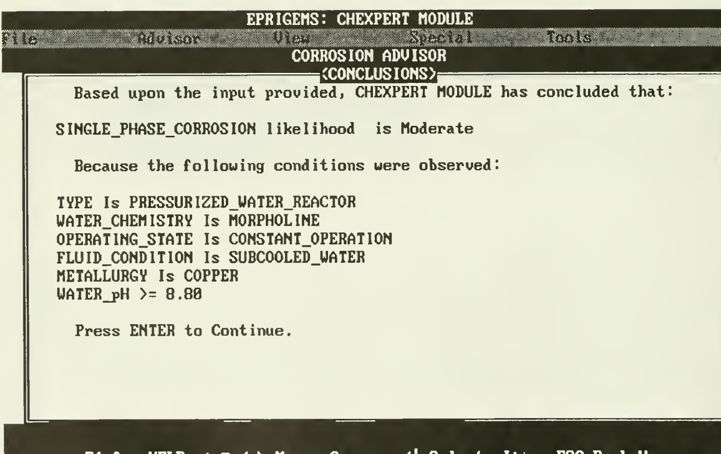


Figure 8: CHEXPERT Evaluation of Corrosion Susceptibility

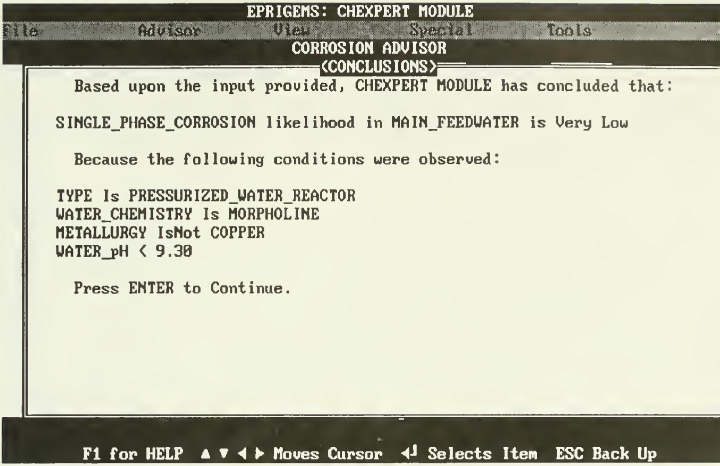


Figure 9: CHEXPERT Situation-Specific Evaluation

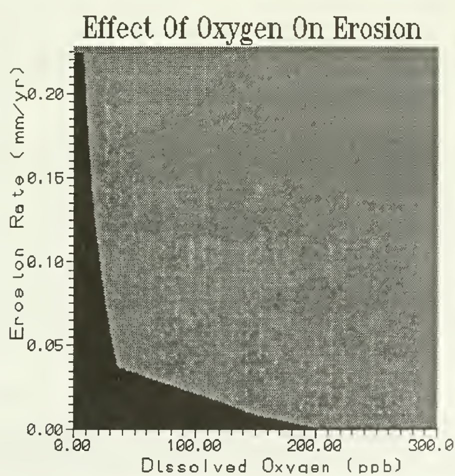


Figure 10: CHEXPERT Quantitative Analysis Display

Cavitation Corrosion Advisor

The Cavitation Corrosion Advisor sub-module is structurally similar to the single- and two-phase modules described above, but contains a completely different rule set aimed at evaluating the potential for the occurrence of cavitation rather than the potential for the occurrence of corrosion. The basic assumption of this module is that the potential for corrosion given that cavitation is occurring is very high. This sub-module treats the same sub-systems as the single-phase module, but considers only those locations (e.g., pump suctions, valve outlets, etc.) where flow cavitation is likely to occur. In addition, since cavitation-assisted corrosion is not specifically treated by the CHEC/CHECMATE analysis programs, susceptibility and situation-specific evaluations are based more upon qualitative rather than quantitative evaluations than either the single- or the two-phase modules.

MIC Advisor

The MIC (Microbially-Induced Corrosion) Advisor sub-module is similar in form to the previous modules but, in many ways, very different in function. No accepted technology exists to support quantitative analysis of MIC, and the underlying physical processes that govern it are completely different from those that govern flow-assisted corrosion. Therefore, the MIC Advisor module relies entirely upon qualitative analysis for both the Susceptibility selection and the Situation-Specific Evaluation selection. In addition, unlike flow-assisted corrosion, the process of evaluating susceptibility to MIC has almost little in common with the process of determining the existence of MIC, so these selections access MIC-specific sub-modules of the Corrosion Advisor diagnostic knowledge base which are totally separate from each other.

The information required to process the MIC susceptibility knowledge base module is similar to that required for flow-assisted corrosion (i.e., metallurgy, operating conditions, etc.), as is the way in which the knowledge base is processed (i.e., goal-driven). However, with MIC, evaluation of susceptibility is a purely qualitative process in which the sub-system is assumed to be susceptible unless it is determined to be impossible. Therefore, while the flow-assisted corrosion modules attempt to compare the supplied information to the results of detailed quantitative analyses to determine susceptibility, the MIC module is limited to a few qualitative tests to determine if MIC is a plausible mechanism in the selected sub-system. The MIC module is thus limited to a two-category susceptibility rating (Possible, Impossible) based primarily upon considerations of water chemistry, metallurgy and operating characteristics of the sub-system.

The MIC situation-specific evaluation knowledge base module is completely different from the flow-assisted corrosion module in that it uses a goal-driven approach to determining the existence of MIC in the selected sub-system. It is also completely different from the MIC susceptibility module in that this module assumes that MIC is the least likely corrosion mechanism in any plant sub-system and that MIC should be assumed only if none of the other mechanisms are plausible. Therefore, in order to establish the existence of MIC, this module evaluates the relative susceptibility of the selected sub-system to each of the other four corrosion mechanisms, then establishes the existence of MIC if corrosion is observed but the susceptibility rating of all other mechanisms is Low. Once the existence of MIC is established, the module then uses a data-driven approach based upon strictly qualitative observations (e.g., size and color of the corroded area, etc.) to determine the type and severity of MIC in the selected sub-system.

IGSCC Advisor

The IGSCC (Inter-Granular Stress Corrosion Cracking) Advisor is structurally similar to the MIC Advisor described above, but somewhat more detailed and quantitative in its treatment of both system susceptibility and situation-specific evaluations. Unlike MIC, IGSCC is a mechanism whose underlying physics are understood and quantifiable based upon readily available metallurgical and chemical information. However, unlike flow-assisted corrosion, IGSCC has not been the subject of extensive quantitative analysis using EPRI analytical programs, so this module remains restricted to mostly qualitative evaluations for both the Susceptibility and the Situation-Specific Evaluation selections.

CHEXPERT KNOWLEDGE BASE

The CHEXPERT Corrosion Advisor diagnostic knowledge base, as discussed briefly in the preceeding section, is a modular knowledge base. The topmost level of the knowledge base hierarchy is the generic diagnostic knowledge base, which is accessed from the diagnostic option of the Corrosion Advisor level menu bar. The purpose of this knowledge base module is to assist the user in determining which of the five corrosion mechanisms treated by CHEXPERT is the most likely mechanism in a particular situation so that he may select this mechanism for more detailed evaluation. This determination is made by first volunteering the information that a corrosion problem exists, then proceeding in a data-driven (forward chaining) mode to determine which of the five mechanisms is the most likely cause of that corrosion. The inputs to this module consist of basic information about the chemistry, metallurgy and operating history of the particular plant in question, supplemented as required by more specific information such as the plant subsystem or piping run of interest. This information is then tested against knowledge base rules which relate various combinations of corrosion "symptoms" to each corrosion mechanism in probabilistic fashion according to the uncertainty analysis treatment described later in this technical paper. The output of this evaluation is a ranking of likely corrosion mechanisms, with the most likely mechanism automatically selected for further evaluation.

The second level of the knowledge base hierarchy consists of a collection of parallel knowledge base modules which perform specific evaluations of sub-system susceptibility to each corrosion mechanism and situation-specific evaluations of the existence and severity of each mechanism. For single-phase, two-phase, cavitation and IGSCC, the sub-system susceptibility knowledge base module performs a goal-driven (backward chaining) evaluation to determine whether or not a particular sub-system is susceptible to that form of corrosion. This evaluation utilizes the same uncertainty analysis treatment as the generic diagnostic knowledge base described above, so the output of this evaluation is a quantitative susceptibility ranking which is converted to a qualitative (i.e., High, Moderate, Low) ranking for display to the user. As described earlier, the susceptibility module for MIC performs a completely deterministic evaluation which does not utilize uncertainty treatment.

The situation-specific evaluation module for all mechanisms except MIC is essentially a continuation of the generic diagnostic module. It performs a data-driven, probabilistic evaluation of the likelihood that the particular form of corrosion exists. The output of this module is a quantitative assessment of this likelihood, together with specific recommendations for preventive or corrective actions. For MIC, the output is the same but the evaluation method is goal-driven based upon the assumption that MIC exists only if no other mechanism is plausible.

For purposes of operating efficiency and ease of maintenance, each of the knowledge base modules described above is stored as a separate knowledge base file that is loaded as needed for processing by the NEXPERT inference engine.

Rule Structure in the CHEXPERT Knowledge Base

The NEXPERT inference engine is a production rule-based expert system which incorporates selected object-oriented programming techniques. Specifically, NEXPERT treats the conclusion of each rule as a boolean (i.e., True/False) object and constrains the conditions of each rule to evaluations of the value of properties of specific objects. In the CHEXPERT knowledge base modules, for the purpose of simplicity and to support the requirements of the uncertainty analysis module described below, all rules contained in a particular module reference properties of a single object whose "name" is a six-character abbreviation of the particular plant under consideration. For example, a rule which tests for the existence of IGSCC at a plant named ABCDEF might read:

If ABCDEF.METAL_CONTENT IS 304SS, then IGSCC_IS_LIKELY

In the above rule, ABCDEF is the object, METAL_CONTENT is its property and IGSCC_IS_LIKELY is the boolean conclusion. Each rule of this type relates a single "symptom" to a specific conclusion, and the sum of the rules with a given conclusion represents the entire "body of evidence" in favor of that conclusion. The methodology used to quantify this "evidence" is described below.

Uncertainty Handling in the CHEXPERT Knowledge Base

A common and serious limitation of many rule-based expert systems is that the rules can only be processed in a purely deterministic manner. For example, the rule:

if A then B

is interpreted as:

if I know that "A" is true, then I know that "B" is true.

However, in power plant applications (and most other "real world" applications) one is never really certain about either the actual value of "A", or the relationship between "A" and "B". In these situations, the above rule should actually be interpreted as:

if I observe that "A" is true, then "B" might also be true.

Although a small number of expert system shell programs incorporate a provision for treating uncertainty, none (including NEXPERT) treat uncertainty in a mathematically rigorous manner that is consistent with the requirements of a power plant diagnostic application. Required features of an uncertainty model for power plant performance diagnosis include:

1. The model must be capable of treating measurement uncertainty (i.e., if I observe that "A" is true, how certain am I that "A" is actually true) and relational uncertainty (i.e., if I know for certain that "A" is true, how certain am I that "B" is true) as separate components of an overall rule uncertainty. This separation of uncertainty components is necessary because measurement uncertainty may vary significantly from instrument to instrument and plant to plant while relational uncertainty remains relatively constant.

2. The model must be capable of treating uncertainty in a form that is conveniently supplied by the domain expert. For example, experience has shown that performance engineering experts find it difficult to quantify the relational uncertainty of the rule expressed above (i.e., if I observe symptom "A", what is the likelihood that it is caused by malfunction "B") because symptom "A" may be a condition common to several malfunctions. However, experts feel much more comfortable in quantifying the uncertainty of the converse relationship (i.e., given that malfunction "B" is true, how certain am I that I should observe symptom "A").

3. The model must be able treat a situation of partial ignorance about a particular measurement or relationship. For example, given the following two rules:

if A then B

if not A then C

If "A" is observed to be true with 80% certainty, one should not automatically assume that "A" is false with 20% certainty because this assumption "creates" evidence in favor of conclusion "C" that may not really exist. Unless there exists some "reason to believe" that "A" is actually false, this remaining 20% certainty should be treated as ignorance about the value of "A".

CHEXPERT addresses all of the above requirements by evaluating rule uncertainty using the Dempster-Shafer Theory of Uncertain Evidence[9]. Dempster-Shafer Theory is ideally suited to power plant diagnostic applications because:

1. It was developed specifically to support an "evidential reasoning" process in which a conclusion is reached based upon the accumulation of supporting evidence rather than an "all-or-nothing" deterministic approach. Dempster-Shafer Theory is therefore completely consistent with the structure of the CHEXPERT knowledge base.

2. It explicitly treats the concept of partial ignorance through use of a dual-value measure of certainty (i.e., certainty about the actual state of a particular parameter is expressed as two values; the first representing the degree of certainty that the observed state is true and the second representing the degree of certainty that the observed state is false). Since the two certainty values are not required to sum to unity, any remaining "unassigned" certainty is attributed to ignorance.

3. It provides an expression for combining uncertainties (Dempster's Rule) that is a natural extension of Bayesian Probability Theory and has been demonstrated to be mathematically rigorous[10]. Dempster's Rule is also sufficiently straightforward to allow it to be manipulated to suit the needs of a particular application.

4. It can be implemented in the NEXPERT expert system shell program through external routines that are executed after successful firing of individual production rules.

Dempster-Shafer Theory represents the current state-of-the-art in uncertainty analysis. Its use in CHEXPERT represents a significant improvement over deterministic or simple Bayesian approaches.

SUMMARY AND CONCLUSIONS

The design and implementation of CHEXPERT, an expert system for corrosion evaluation, have been described. This shows how expert system technology can provide the user with the capability to:

1. Understand the various corrosion mechanisms;

2. Recognize if a corrosion problem exists in his plant;

3. Identify the possible corrosion mechanisms responsible for the problem;

4. Identify the possible remedies for the problem and how to implement them. These include practical techniques, EPRI's analytical tools, reports and experts.

It is expected that such a system which combines both educational and diagnostic features will prove valuable to the plant engineer. Furthermore, in conjunction with predictive tools developed by EPRI, the plant engineer can plan and implement a sound, long-term imspection program based on state of the art knowledge to prevent catastrophic failures.

CHEXPERT will be further refined as user feedback becomes available. These refinements may include more detailed tutorials or diagnostics, additional references and additional corrosion mechanisms.

REFERENCES

1. CHEC Computer Program User's Manual. Palo Alto, CA: Electric Power Research Institute, February, 1988.

2. CHECMATE Computer Program User's Manual. Palo Alto, CA: Electric Power Research Institute, April, 1989.

3. EASE+ TOOLS Reference Manual. Belmont, CA: Expert-EASE Systems Inc., November, 1988.

4. NEXPERT OBJECT Fundamentals. Palo Alto, CA: Neuron Data Inc., 1986.

5. R. Jones et. al. "Single-Phase Erosion-Corrosion of Carbon Steel Piping." Palo Alto CA: Electric Power Research Institute, February, 1987.

6. MICPro User's Manual. San Jose, CA: Structural Integrity Associates, 1988.

7. G. Licinia. "A Review of Microbial Induced Corrosion in Nuclear Power Plant Systems." St. Louis, MO: NACE Publications, 1988.

8. IGSCC-EX User Manual. San Jose CA: Structural Integrity Associates, 1988.

9. G. Shafer. "Mathematical Theory of Evidence." Princeton, NJ: Princeton University Press, 1976.

10. G. Liu. "Causal and Plausible Reasoning in Expert Systems." Berkeley, CA: University of California at Berkeley, 1987.

An Expert System for Microbiologically Influenced Corrosion

C. E. CARNEY and G. J. LICINA
Structural Integrity Associates
San Jose, California, USA

ABSTRACT

Microbiologically Influenced Corrosion (MIC) is a damage mechanism that can cause serious degradation of service water system components. MIC can be particularly insidious since damage can occur very quickly, even in environments otherwise resistant to corrosion. Plant operations or maintenance personnel or system engineers typically do not have sufficient expertise to predict when and where MIC may occur or what methods of treatment are effective. An expert system (MICPro) has been devised which provides a tool for utilities to predict where MIC will occur, which systems or components are most susceptible, how operating parameters may affect vulnerability, and how to implement corrective and preventative measures. The system is designed to be simple to use: required inputs are common system parmeters and results are presented as numbers from 1 to 10 indicating the likelihood of damage due to the given input. The structure and operation of the system is described, and future refinements are discussed.

BACKGROUND

Microbiologically Influenced Corrosion (MIC) involves the interaction between biological activity and the electrochemical process of corrosion. MIC is one of the few corrosion mechanisms that is operative at low temperatures and one of the only mechanisms that affects components under stagnant conditions. MIC can afflict essentially all systems of a nuclear power plant and can seriously degrade the life of components in very short times. (For example, through–wall pitting of stainless steel piping systems left in contact with potable water — used for hydrostatic testing — for just one or two months can proceed at an average rate of penetration on the order of inches per year). MIC may be the prime contributor to the degradation of systems or components that are either: (a) in contact with untreated water for any significant period of time (such as plant construction or extended lay–up), or (b) that are typically maintained in a standby mode, or (c) that experience long periods of stagnation or of very low flow. Many components which fit these descriptions are virtually inaccessible for repair. Many are safety related systems or support safety systems. The flow capabilities of some lines may also be affected as massive quantities of corrosion products are deposited resulting in serious restrictions to flow capabilities including complete blockage of the line.

The loss of flow in safety related systems, or even in systems that provide cooling water to safety related equipment, provides a serious concern to the plant owner. Concerns with MIC have prompted a Nuclear Regulatory Commission Inspection and Enforcement Bulletin [1] and a Significant Events Report from the Institute of Nuclear Power Operations [2]. Utilities have devoted increasing attention to problems related to raw water service including a number of instances where pipe has been replaced, often with extremely expensive stainless grades, in an attempt to alleviate MIC–related operational difficulties. The Electric Power Research Institute and individual utilities have devoted an increasing level of attention to the breadth of service water system problems, with an emphasis on corrosion problems including MIC.

Further, there is no simple solution to problems of MIC. The application of corrective actions to situations where MIC is suspected rely extremely heavily upon a proper diagnosis. A correct diagnosis is of particular importance since treatments for MIC are not only expensive, but improper or unnecessary application of biocide can actually induce new corrosion mechanisms or aggravate existing corrosion conditions resulting from other sources. Guidelines and philosophy for obtaining a correct diagnosis have been emphasized in the EPRI and NACE documents on MIC [3–6]. For instance, the MIC sourcebook [3] recommends that a thorough diagnostic procedure be followed attempting to prove that the corrosion is due to causes other than biological activity — "MIC should be concluded as the cause of, or a contributor to, the observed attack only if the situation cannot be explained by other means."

Although the existence of microbiologically influenced corrosion is well established, the bulk of the publications on the prevention, detection, and treatment of MIC remain in the R&D domain. NACE and EPRI have recently published guidelines on the prediction, diagnosis, and mitigation of MIC [3,5,6]. However, the actual application of those guidelines to particular plant situations still generally requires a more detailed understanding of the mechanisms of, and contributors to, MIC (i.e., more expertise) than most personnel concerned with plant operations would care to obtain. (Also, work in the subject is very active, as experiments and field data periodically uncover new problems and corrosion mechanisms. Keeping up with the latest developments can consume more time for plant personnel than they have available for such efforts.)

To fully protect their service water systems, utilities need methods for prediction of where MIC may occur, which systems are most susceptible, how operational parameters may affect vulnerability of components, and how to treat existing MIC problems and prevent future ones. Such methods may further require the ability to examine components that have failed due to corrosion and to determine what mechanisms (MIC, non–MIC) were involved in the failure. Since operations or maintenance personnel or system engineers typically do not have the revevant expertise to make such predictions or judgements themselves, (and cannot reasonably obtain it) the use of an expert system, with a knowledge base developed from research experiences and from the expertise of others permits a rapid, interactive method for utility personnel to access the expert knowledge and apply it to their plant systems.

MICPro is an expert system developed to address these needs. The **MICPro** knowledge base contains the information from the EPRI MIC sourcebook [3] plus additional information that has been collected since the sourcebook was issued in 1988. This expert system was produced by the authors under guidance of EPRI project RP2939–1.

PROGRAM DESIGN

MICPro was developed to provide the system engineering, water chemistry, materials engineering, or maintenance specialist access to the expertise required to predict where MIC might be expected, the relative contributors to attack, and potential methods for mitigation. These target users of the system and their needs defined much of the overall design. The system must be simple to use, or people will not choose to use it. Since these personnel may have no training in biological mechanisms, the system should not use technical language, but should relate MIC directly to operational information. System configuration and operation provide the inputs. Output is a simple set of ratings, on a scale of 1 to 10, reflecting the susceptibility of that system or component to damage by MIC (and also a similar index for corrosion without biological influences). Further, the system should be able to provide intelligent defaults when the user is unsure of some parameters. Help messages should be available to advise the user on input values desired and on interpreting the results.

A further design decision was made to limit the scope of this system. Rather than trying to produce a complete (and therefore more complex) MIC expert, **MICPro** was designed as a simple tool to acheive limited objectives — to predict damage due to MIC in service water systems, and to give guidance in the diagnosis of MIC failures (including an evaluation of abiotic corrosion for comparison). Thus, the full conception of the **MICPro** expert system includes 2 functional units: a predictive advisor to assist with vulnerability predictions and with failure analysis for specific locations in systems where MIC might be anticipated, and a diagnostic advisor that will assist the failure analyst in selecting the type of analytical techniques and physical tests to use to determine whether or not a failure has been influenced by microbiological activity. (At this point in time, only the predictive mode of operation is available — however, this function provides some diagnostic support as well, as detailed below.)

The EPRI–generated expert system SMART (SMall Artificial Reasoning Toolkit) [7] was used as a shell for the system. The SMART shell was chosen for several reasons. First, since the authors were working on an EPRI–sponsored project, this shell was easily available (free) and presented no difficulties of licensing. Second, through work on other projects, the SMART shell was familiar to the authors. Third, SMART is both flexible and extensible, a feature which turned out to be very important in tuning some of the non–standard reasoning approaches used. Finally, SMART suports a user interface based on the EPRIGEMS specification, which provides a standard look and feel that may be familiar to utility personnel using the system.

MICPro's program logic is strongly influenced by the decision to present the evaluation results as a single number (the System Index) that indicates the degree to which MIC (or abiotic corrosion) might be expected for the component or system in question. To determine this Index, **MICPro** first computes several sub–indices, each one reflecting the independant contributions to corrosion due to some operational or system parameter known to be signifigant. (Specifically, material, water chemistry, temperature (and d T), water treatments, and operating flows are used.) The program then combines and weighs the various contributions to determine the overall System Index. The System Indices and sub–indices for material, water, flow, and temperature are given on a 1 to 10 scale where 1 represents extreme resistance to MIC or corrosion and 10 represents extreme susceptibility. (An index of zero is used in unusual cases to indicate an immunity to MIC.) Numerical combining rules were devised and weighted to account for the direct interactions of the key variables. For example, such parameters as the length of stagnant periods, the number of stagnant periods, etc. are compared to the system operating life and assigned indices that describe the contribution of that flow history to MIC susceptibility. Special rules were prepared to account for combinations of factors with unusual results (i.e. the strong corrosive effects of chlorine–based biocides on carbon steels).

Initially, the combining rules for all of the parameters were set to produce a simple multiplicative average, a simple rule that modeled the expert's expectations of combination effects. As development proceeded, special rules and weights were added to account for special combinations of factors and special cases where one or two single factors controlled the corrosion process. Once initial coding was complete, many test cases were run and the results examined closely to fine–tune the rules to yield reasonable System Indices over a wide variety of conditions (i.e., material, water chemistries, flow, temperature, and treatment). This method of closer approximations proved very effective: the final version of the combining rules was tested using virtually all of the cases described in the MIC Sourcebook [3] and gave final ratings that were always consistant with the actual corrosion present.

Constructing the combining rules represented a deviation from the normal types of reasoning used to build an expert system's inference engine. In its issued form, the SMART shell was unable to handle the numeric inputs, combining rules, and outputs. Some modifications were required to the shell to permit this more quantitative approach to the analysis. However, the autors believe that this effort was justified, since the end result is a final report that is clear and informative even to users with no biological background, the reasoning follows the intuitive judgements of experts, and the conclusions are accurate.

The Predictive Advisor of MICPro performs the analysis using a combination of forward chaining logic and direct calculation. Once the input forms are completed, forward chaining proceeds to set default values and note special cases in factor combination. Any logical conclusions that may be of interest to the user are saved for the report. Then, each of the sub–indices is computed, and these are in turn combined to produce the two System Indices.

PROGRAM OPERATION

The predictive mode of MICPro permits assessments of the relative susceptibility of systems and locations within systems to MIC based upon the materials of construction, the operating history, water chemistry, and water treatment. A session with the MICPro Advisor proceeds thru three stages: Input, evaluation, and reporting results.

During the input stage, values for all of the key variables are input by the user at several input forms. (see Figures 1 thru 6) Default values will be assigned intelligently by the advisor if a required data field is not filled. The MICPro Predictive Advisor then processes the given input data, computing the various sub–indices and searching its knowledge base for any special–case rules that apply. MICPro then gives a report that includes the System Index and the sub–indices for the specified system/component. An example of this report is included as Figure 7.

Evaluations of susceptibility to both MIC and corrosion without biological influences are given in the Predictive Advisor's report, primarily to alert the user that all corrosion in untreated water is not necessarily MIC. Many natural waters which are rich in bacteria that promote corrosion are also very corrosive without any biological enhancement. The corrosion index in the report is provided to alert the user that even for waters where the susceptibility to MIC may be high, the susceptibility to corrosion in the same water, even if that water were sterile, would still be high. In such cases, differentiation between MIC and corrosion due to the water chemistry and component operating conditions requires additional investigation.

Several report options are included to permit the results of the analysis to be reviewed (on the computer monitor), saved to a disk for future editing, or printed. The report consists of all of the information included on the input forms, plus a summary table of the system indices for MIC and for (abiotic) corrosion, along with a list of conclusions reached in the evaluation that serves to explain how the numerical values were determined.

Help messages are provided at all levels to assist with data entry and to explain the importance of a particular value to the analysis. For many inputs a list of options is offered (e.g., materials of construction, product forms, or water sources) so that the user may select an item from the list rather than typing its name in. The user is also given the option of saving the input data on on a restart file such that any inputs may be saved from one run to the next, even if the computer is turned off.

The final reports (shown in Figure 7) provide information that may be used in a number of ways. First, the user can determine which corrosion mechanisms, if any, will be applicable within his systems. The computed System Indices for MIC and abiotic corrosion are listed along with a description of the relative susceptibility (Low, Moderate, High, Very High, etc.) in the first report. If either or both indices are greater than approximately seven, the system would be expected to experience corrosion (from microbiological influences or from more "conventional" sources). If both indices are less than five, little corrosion would be expected. If one of the indices is high (> 7) and the difference between the two indices is

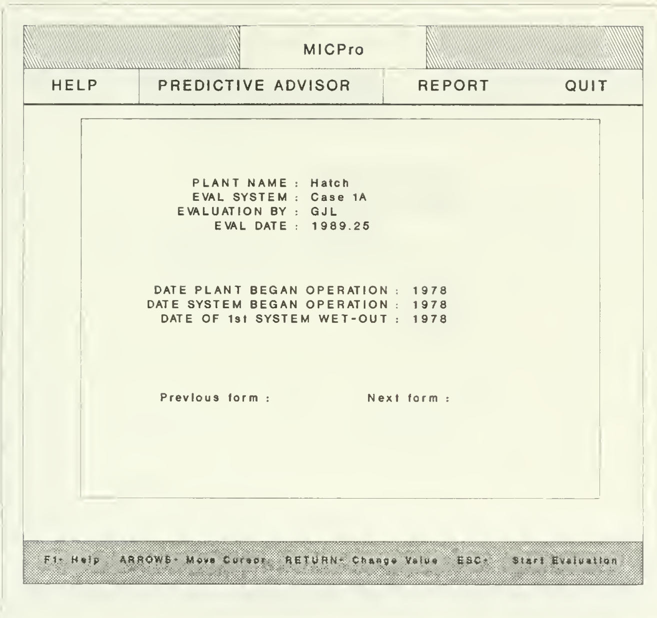


Figure 1. MICPro General Input Screen

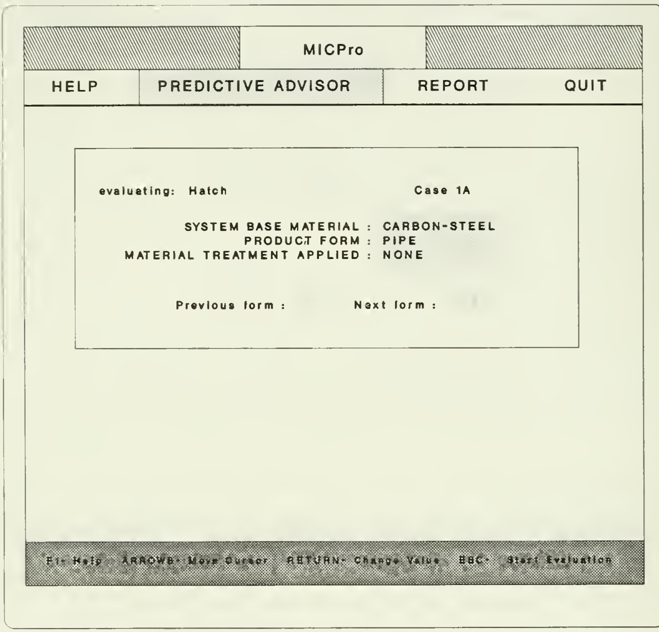


Figure 2. Materials Input Screen

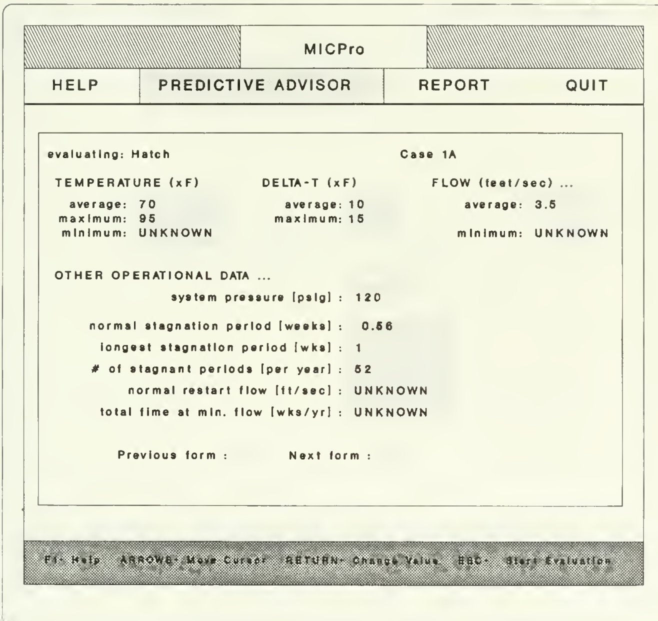


Figure 3. Operational Data Input Screen

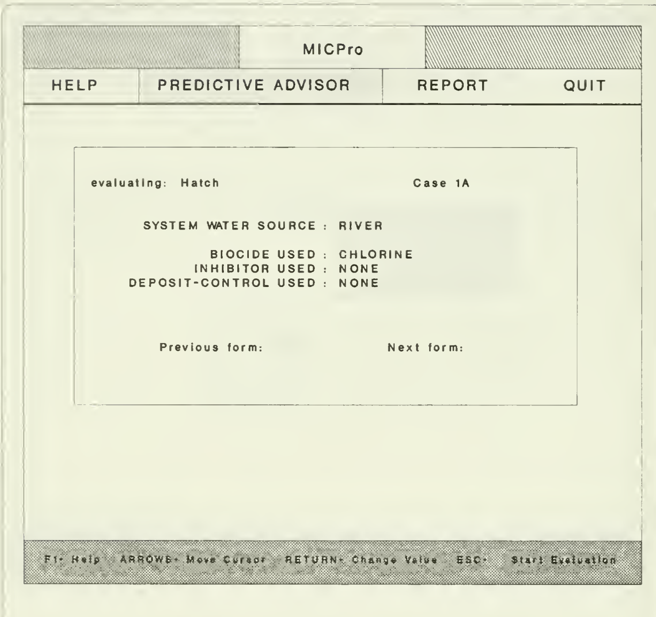


Figure 4. Water Source Input Screen

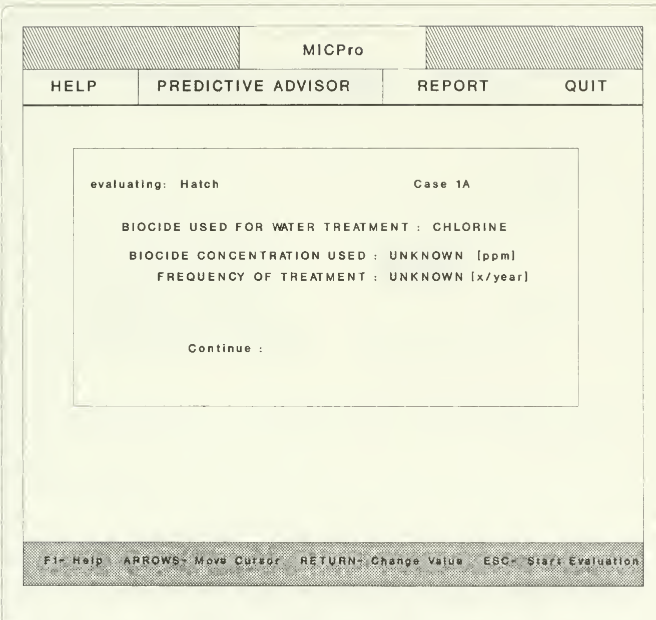


Figure 5. Water Treatment Input Screen

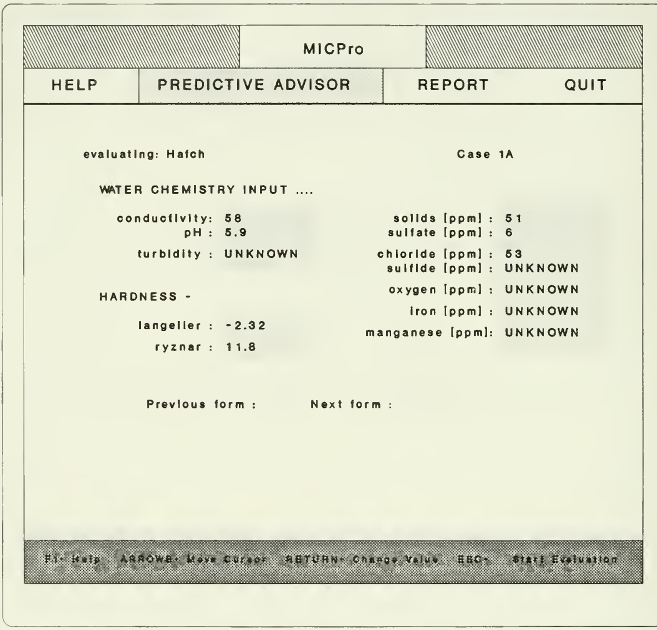


Figure 6. Water Chemistry Input Screen

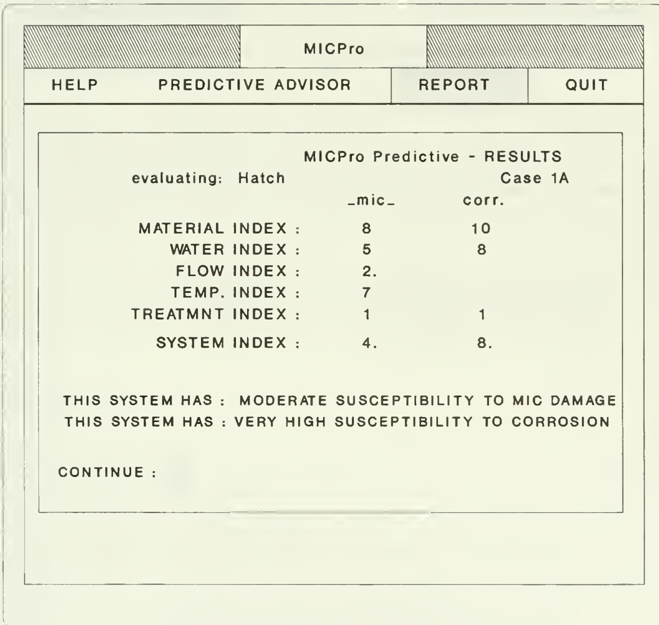


Figure 7. MICPro Results Screen #1 - Susceptibility Report

MICPro

HELP | PREDICTIVE ADVISOR | REPORT | QUIT

THE CONCLUSIONS REACHED IN THIS SESSION ARE:

WARNING: DEFAULT VALUE USED FOR BIOCIDE CONCENTRATION
WARNING: DEFAULT VALUE USED FOR SULFIDE CONCENTRATION
WARNING: DEFAULT VALUE USED FOR TIME AT MINIMUM FLOW
WARNING: DEFAULT VALUE USED FOR MINIMUM SYSTEM FLOW

--> WATER TREATMENT INEFFECTIVE AGAINST MIC
--> SOFT WATER CAN BE AGRESSIVE TO CARBON STEEL
--> OPERATING TEMPERATURE PROMOTES MIC
--> CARBON STEEL IS SUSCEPTIBLE TO MIC

Figure 7. (cont) MICPro Results Screen #2 - Advisor's Conclusions

more than two units (e.g., system index for MIC 7; system index for corrosion 2), corrosion would be expected with the likely source being either MIC (for the values cited above) or the aqueous environment depending upon which index is higher.

Different locations within a system may also be evaluated by simply modifying the inputs to reflect the temperature, flow, biocide concentration or other conditions at that location. Applied in this manner, **MICPro** can be used to pinpoint the most likely vulnerable locations within a system. These locations may be selected for further examination or selected as the best locations for sidestreams containing corrosion coupons, electrochemical probes, or other monitoring and prevention methods.

The sub–indices also provide insight into the relative contributions of material, water chemistry, operating conditions (flow), temperature, and water treatment. A high value for one or more of these sub–indices indicates that that parameter (or parameters) is (are) controlling and presents the most likely candidate for a mitigation treatment. The converse will also be true. That is, the sensitivity to MIC or abiotic corrosion to candidate mitigation measures may be evaluated by simply changing the inputs to reflect the candidate treatment, re–running the analysis, and examining the effect on both the system indices and the various sub–indices.

FUTURE REFINEMENTS

The primary source of information for **MICPro** is the Sourcebook for Microbiologically Influenced Corrosion [3] which is a review of MIC in nuclear power plants; not a detailed tome on corrosion. While this initial version of **MICPro** provides separate indicators to predict the susceptibility to microbiologically influenced corrosion and corrosion due to non–biological factors, the model for evaluating abiotic corrosion is admittedly simplistic. The handling of various water treatments, particularly corrosion inhibitors and deposit control agents, is also very crude. A refinement to the expert system planned for the near future is the incorporation of more sophisticated methods for prediction of abiotic corrosion and handling of typical water treatments. This step will require the debriefing of industry experts in these areas. Preliminary contacts and a course of action have been outlined.

The corrosion and MIC susceptibility evaluations utilize only a few water chemistry inputs. Greater sophistication of the predictive models will be based upon consideration of more details of the water chemistry including the capability for additional calculations of important parameters (e.g., hardness indices).

Only the two most commonly used mitigation measures (water treatment and materials replacement) are addressed in this version of MICPro. Future work on MICPro will also include alternative mitigation measures such as cathodic protection, water treatment with ultraviolet light, filtration through media of very fine size (on the order of microns), and heat disinfection. Subsequent versions of MICPro will also address cleaning processes in some detail.

SUMMARY

In its present form, MICPro gives the user a tool for making predictions of the susceptibility of systems, or specific locations within those systems, to attack due to MIC. MICPro also provides a simple method for evaluating the likely effectiveness of candidate mitigation measures. Correct diagnosis is extremely important in all cases where MIC may be operative since most treatments to mitigate MIC are expensive. Even more importantly, the consequences of a "false positive" (i.e. , concluding that microbiological effects are influencing corrosion when they actually are not) can actually exacerbate corrosion when the "real" problem is corrosion due to a naturally aggressive water or under–deposit corrosion. A Diagnostic Advisor has been planned for MICPro that will provide guidelines for sampling and assistance in concluding whether microbiological influences were operative in failure analyses where MIC is suspected.

REFERENCES

1. "Microbiologically Influenced Corrosion of Containment Service Water System", IE Information Notice No. 85–30, United States Nuclear Regulatory Commission Office of Inspection and Enforcement, April 19, 1985.

2. "Microbiologically Influenced Corrosion (MIC)", Institute of Nuclear Power Operations Significant Event Report SER 73–84, 1984.

3. G.J. Licina, "Sourcebook for Microbiologically Influenced Corrosion in Nuclear Power Plants." Palo Alto, CA: Electric Power Research Institute, EPRI NP–5580, 1988.

4. D.H. Pope, "A Study of Microbiologically Influenced Corrosion in Nuclear Power Plants and a Practical Guide for Countermeasures." Palo Alto, CA: Electric Power Research Institute, EPRI NP–4582, May 1986.

5. "The Role of Bacteria in the Corrosion of Oil Field Equipment", TPC Publication 3. Houston, TX: NACE, March 1982.

6. "A Practical Manual on Microbiological–Influenced Corrosion." Houston, TX: NACE, to be published.

7. David Cain, "Small Artificial Reasoning Toolkit." Palo Alto, CA: Electric Power Research Institute, EPRI NP 5645M, July, 1988.

Expert System Application for Oyster Creek

H. FU
GPU Nuclear Corporation
One Upper Pond Road
Parsippany, New Jersey 07054, USA

ABSTRACT

Two PC-based expert systems SMARTRODS and ESAO, have been developed to support Oyster Creek start-up at the Oyster Creek Nuclear Generating Station. SMARTRODS is a LISP program coupled with a user interface which is developed using EPRI-SMART. It generates a control rod withdrawal sequence table for reactor start-up based on the given initial and target control rod patterns. It also checks a given sequence table for rod movement which may result in excessive local power peaks. The reactor core power is monitored by neutron detectors located in the reactor core. Oyster Creek Technical Specifications state the minimum number of and location of detectors required for properly monitoring the core power. During start-up, compliance with these technical specifications has to be checked before the reactor power can be increased. ESAO is a rule-based expert system developed to perform this compliance check. Both expert systems will be tested during Oyster Creek Cycle 12 start-up. This paper describes these two expert systems and their usage at Oyster Creek.

INTRODUCTION

Oyster Creek is a Boiling Water Reactor with a rated power of 630 MWe. The replacement power cost for Oyster Creek is approximately half-a-million dollars per day when the reactor is shut down. It is important that the reactor start-up process is safe and without unnecessary delays. The reactor operators and engineers have to ensure that the reactor core power increase is being properly monitored such that fuel integrity is maintained and thermal limits are not exceeded. During the start-up, they have to make quick and accurate decisions to insure adequate instrumentation is available to monitor power increases and that the control rod withdrawal sequence table is providing the anticipated power increase. These require operation experience and following certain rules-of-thumb. Two expert systems, SMARTRODS and ESAO, are, therefore, developed to support Oyster Creek start-up.

SMARTRODS

SMARTRODS, Rule Ordered withDrawal Sequences with SMART user interface, is an expert system to determine control rod withdrawal sequence table from an initial and a target rod pattern, or to check a given control rod withdrawal sequence table to prevent fuel damage.

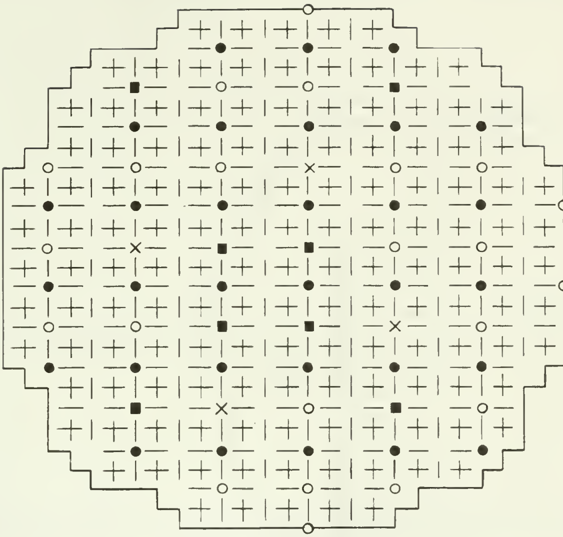
Background

In a nuclear power plant, control rods are used to regulate reactor power. At Oyster Creek, there are 560 fuel assemblies and 137 cruciform control rods, with each control rod inserted between sets of four fuel assemblies. The Oyster Creek core map is shown in Figure 1. At the beginning of reactor start-up, all the control rods are inserted. As the reactor power increases, control rods are withdrawn from the reactor in accordance with the control rod withdrawal sequence table, until the target control rod pattern is reached. Figure 2 depicts a typical response of assembly axial power to control rod withdrawal. It is important that the control rods are withdrawn in such a manner that the local power level does not become excessive, otherwise, the expansion of the fuel pellets due to overheating can cause a fuel rod to rupture and release fission products into the boiling water. The reactor engineers would develop the control rod withdrawal sequence table based on their operating experience prior to the start-up. However, changes in the target rod pattern and control rod withdrawal sequence occur during start-up due to differences in expected power changes to those experienced previously. An expert system for developing and checking withdrawal sequence table would be helpful during start-up; by both saving time and insuring changes can be made quickly and accurately during the start-up.

RODS, expert system for Rule Ordered withDrawal Sequences, was developed in 1983 under a joint research project between MITRE and GPU Nuclear. Mr. J. Reierson of MITRE Corporation was the knowledge engineer, and Mr. R. V. Furia of GPU Nuclear was the domain expert. The rules were developed based on the rules-of-thumb used by Oyster Creek reactor engineers during start-up. RODS was originally written in Franz LISP on a VAX-11/780 computer. Unfortunately, RODS could not be used at Oyster Creek because of the software and hardware requirement. With the IBM PC available, it was decided that RODS should be converted to run on the PC. This was done in 1986, but it was not user friendly since it required the user knew which specific LISP functions to execute in order to initiate the expert system. This made it very difficult for the reactor engineers to use the expert system. With the use of EPRI-SMART, a user interface is added to provide menu for consultation.

System Description

SMARTRODS is RODS with a user interface developed with EPRI-SMART. It runs on IBM PC or compatibles. It is menu-driven with no required user's knowledge of LISP or SMART. When entering the expert system, the user is prompted with the screen shown in Fig.3. The INTRODUCTION option provides general information about SMARTROD and EXIT from the expert system. The INPUT option let user initialize the global data base by entering the data for control rod group location, initial and target rod pattern, and control rod sequence table. When selected, the user is prompted with the screen shown in Figures 4-7. Although a full core map is presented, the user only needs to enter quarter core data, and the system expands it to full core. When OPTIONS is selected, the user can choose (1) to develop control rod withdrawal sequence from all-rods-in to the target rod pattern, (2) to develop control rod withdrawal sequence from an intermediate rod pattern during start-up to the target rod pattern, (3) to check a control rod withdrawal sequence table, or (4) to make step change of a control rod withdrawal sequence table and check the revised table. The user is prompted with the required input for each selection. The input data shown are those stored in the global data base. The user can either



Item	No.
+ Control Rod	137
☐ Fuel Assembly	560

Neutron Monitoring System

●	Local Power Range Monitor (LPRM)	31
x	Source Range Monitor (SRM)	4
■	Intermediate Range Monitor (IRM)	8
o	Spare Penetrations	22

Figure 1

Oyster Creek Core Map

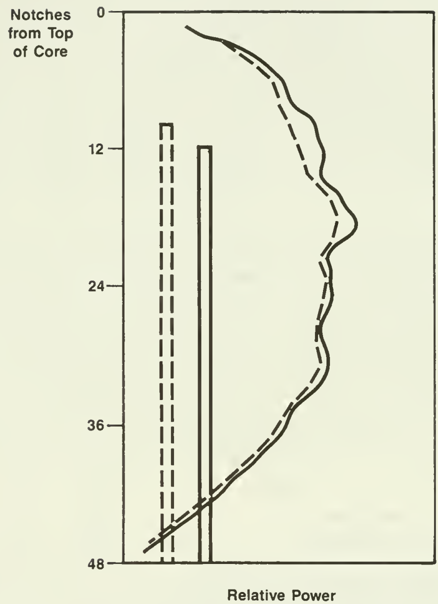


Figure 2

Assembly Axial Power Response to Control Rod Withdrawal

```
                          GPUN SMART-RODS
  INTRODUCTION     INPUT        OPTIONS        OUTPUT

<F1> HELP <ENTER> RUN OPTION <ESC> EXIT OPTION
```

Figure 3
SMARTRODS Menu

```
GROUP #:1.1 2.1 3.1 4.1 5.1 6.1 6.2 7.1 7.2 7.3 8.1 9.1 9.2 9.3 9.4 10.1 10.2

                              GROUP MAP
                    -5.1 -2.1 -6.1 -2.1 -5.1
          -3.1 -6.2 -4.1 -7.1 -3.1 -7.1 -4.1 -6.2 -3.1
     -3.1 -9.1 -2.1 10.1 -1.1 -9.2 -1.1 10.1 -2.1 -9.1 -3.1
     -6.2 -4.1 -7.3 -3.1 -8.1 -4.1 -8.1 -3.1 -7.3 -4.1 -6.2
-5.1 -2.1 10.1 -1.1 -9.3 -2.1 10.2 -2.1 -9.3 -1.1 10.1 -2.1 -5.1
-4.1 -7.1 -3.1 -8.1 -4.1 -7.2 -3.1 -7.2 -4.1 -8.1 -3.1 -7.1 -4.1
-6.1 -1.1 -9.2 -2.1 10.2 -1.1 -9.4 -1.1 10.2 -2.1 -9.2 -1.1 -6.1
-4.1 -7.1 -3.1 -8.1 -4.1 -7.2 -3.1 -7.2 -4.1 -8.1 -3.1 -7.1 -4.1
-5.1 -2.1 10.1 -1.1 -9.3 -2.1 10.2 -2.1 -9.3 -1.1 10.1 -2.1 -5.1
     -6.2 -4.1 -7.3 -3.1 -8.1 -4.1 -8.1 -3.1 -7.3 -4.1 -6.2
     -3.1 -9.1 -2.1 10.1 -1.1 -9.2 -1.1 10.1 -2.1 -9.1 -3.1
          -3.1 -6.2 -4.1 -7.1 -3.1 -7.1 -4.1 -6.2 -3.1
                    -5.1 -2.1 -6.1 -2.1 -5.1


NOTE:Only need to enter data for quarter core
     Hit Esc and reenter to check full core data
```

Figure 4
Input Screen for Control Rod Group Map

```
Enter the withdrawal sequence step corresponding to this rod pattern 26

                          INITIAL ROD PATTERN

                        48 48 48 48 48
                  48 48 48 48 48 48 48 48 48
               48 00 48 00 48 00 48 00 48 00 48
               48 48 48 48 48 48 48 48 48 48 48
            48 48 00 48 00 48 00 48 00 48 00 48 48
            48 48 48 48 48 48 48 48 48 48 48 48 48
            48 48 00 48 00 48 00 48 00 48 00 48 48
            48 48 48 48 48 48 48 48 48 48 48 48 48
            48 48 00 48 00 48 00 48 00 48 00 48 48
               48 48 48 48 48 48 48 48 48 48 48
               48 00 48 00 48 00 48 00 48 00 48
                  48 48 48 48 48 48 48 48 48
                        48 48 48 48 48

NOTE:Only need to enter data for quarter core
     Hit Esc and reenter to check full core data
```

Figure 5
Input Screen for Initial Rod Pattern

```
                          TARGET ROD PATTERN


                        48 48 48 48 48
                  48 48 48 48 48 48 48 48 48
               48 28 48 14 48 -6 48 14 48 28 48
               48 48 48 48 48 48 48 48 48 48 48
            48 48 14 48 00 48 22 48 00 48 14 48 48
            48 48 48 48 48 48 48 48 48 48 48 48 48
            48 48 -6 48 22 48 16 48 22 48 -6 48 48
            48 48 48 48 48 48 48 48 48 48 48 48 48
            48 48 14 48 00 48 22 48 00 48 14 48 48
               48 48 48 48 48 48 48 48 48 48 48
               48 28 48 14 48 -6 48 14 48 28 48
                  48 48 48 48 48 48 48 48 48
                        48 48 48 48 48

NOTE:Only need to enter data for quarter core
     Hit Esc and reenter to check full core data
```

Figure 6
Input Screen for Target Rod Pattern

CR WITHDRAWAL SEQUENCE TABLE

BEGINNING STEP NUMBER:32 LAST STEP NUMBER:47

GROUP-1.1	48															
GROUP-2.1	48															
GROUP-3.1	48															
GROUP-4.1	48															
GROUP-5.1	48															
GROUP-6.1	48															
GROUP-6.2	48															
GROUP-7.1	48															
GROUP-7.2	48															
GROUP-7.3	44	48														
GROUP-8.1	32	36	40	42	44	48										
GROUP-9.1	0	0	4	8	12	16	20	24	26	28	30	30	30	30	32	34
GROUP-9.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
GROUP-9.3	0	0	0	0	0	0	0	0	0	0	0	0	4	6	8	10
GROUP-9.4	0	0	0	0	0	0	0	0	0	0	4	6	8	10	12	14
GROUP-10.1	0	0	0	0	0	4	8	10	12	14	16	16	16	16	18	20
GROUP-10.2	8	12	16	20	24	26	28	30	32	34	34	34	34	34	36	38

Hit ESC key to continue

Figure 7
Input Screen for Control Rod Withdrawal Sequence

change the input data or hit Esc key to continue. It is frequently necessary to alter a withdrawal sequence during start-up, the CHANGE-STEP option allows user to make three single step value changes and check the revised table. The RESULTS option is the same as OPTIONS except it writes all the results to a data file, instead of the monitor. Later these output files can be printed or saved for permanent record.

ESAO

ESAO, Expert System for APRM Operability, is a ruled-base expert system for determining the operability of Averaged Power Range Monitors (APRM) and check the related Technical Specification compliance.

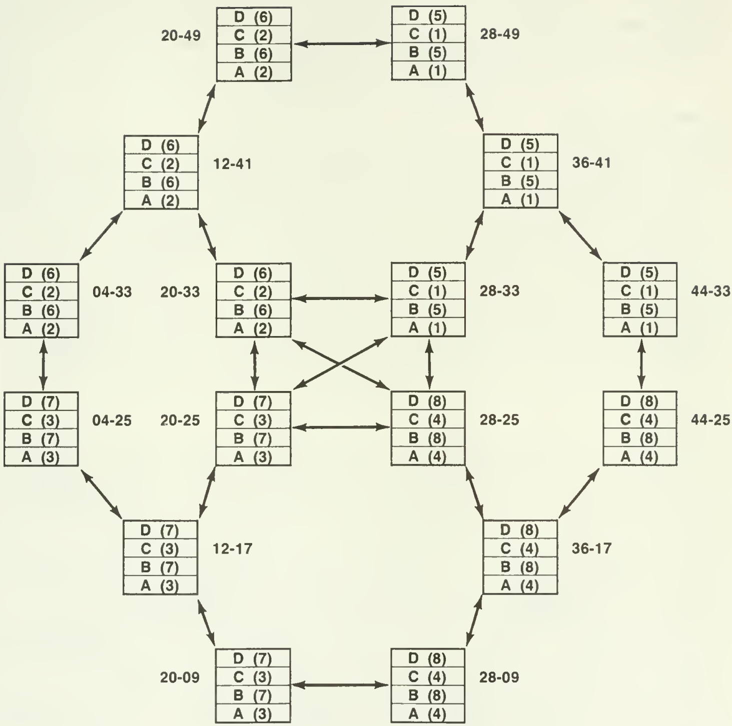
Background

Oyster Creek has three levels of neutron detectors: source range monitors for very low power; intermediate range monitors for low power; and power range monitors for low to high power. The power range monitors measure the power at each detector location and provide input to the average power range monitor (APRM). There are 16 local power range monitoring (LPRM) strings distributed uniformly about the reactor core. Each LPRM string contains four detector located at fixed axial locations. Signals from the 64 detectors are fed into eight averaging circuits (APRMs) covering each quadrant of the reactor core as shown in Figure 8.

Oyster Creek Technical Specification states the following for determining operability of protective instrumentation:

- 3.1.A. One APRM in each operable trip system may be bypassed or inoperable provided the requirements of specification 3.1.C and 3.10.C are satisfied. Two APRM's in the same quadrant shall not be concurrently bypassed except as noted below or permitted by note.
- 3.1.B.1. Failure of four chambers assigned to any one APRM shall make the APRM inoperable.
- 3.1.B.2. Failure of two chambers assigned to any one radial core location in any one APRM shall make that APRM inoperable.
- 3.1.C.1. Any two LPRM assemblies which are input to the APRM system and are separated in distance by less than three times the control rod pitch may not contain a combination of more than three inoperable detectors out of the four detectors located in either the A and B, or the C and D levels.

It is important that these specifications be met during reactor operation to ensure that local reactor power has been properly monitored. During reactor start-up power level is monitored from the source range to intermediate range to the power range. Prior to switching from the intermediate range into the power range the reactor operator must insure there are an adequate number of local power range detectors available to meet the above specification. A detector can be failed or if it is reading downscale it must be bypassed. Only



A, B, C, D: LPRM detector level

(1), (2), (3), (4), (5), (6), (7), (8): APRM channel

XX-YY: LPRM location

⟷: LPRM within 3 control rod pitches

Figure 8

Oyster Creek APRM Configurations

a limited number of detectors can be failed or bypassed. Before the operator can switch to the power range monitors, he needs to know if the above conditions have been completed. Otherwise, the operator must wait for the reading to come on scale prior to switching, thus delaying the start-up This is sometimes accomplished by the reactor engineer adjusting the control rod withdrawal sequence. Therefore, a quick and accurate determination of the technical specification compliance is desirable.

System Description

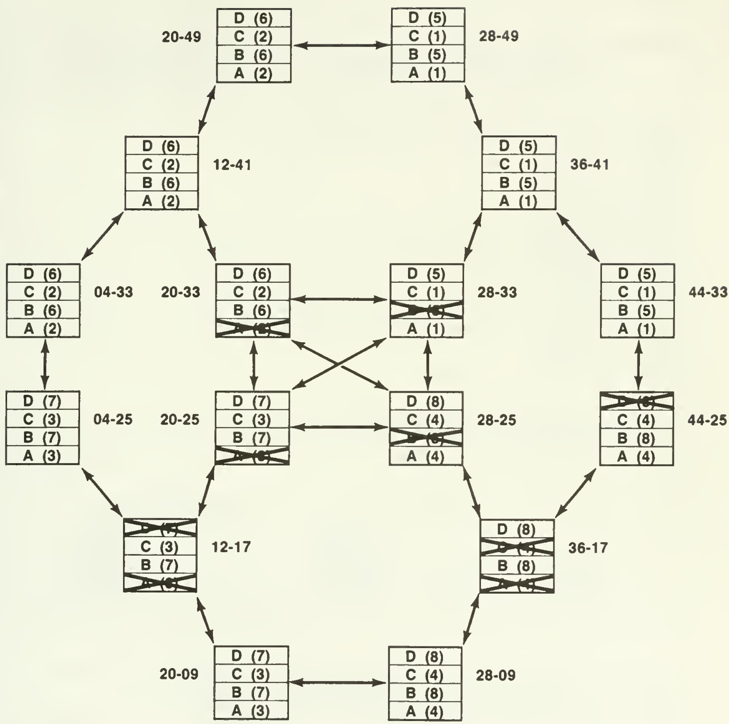
ESAO is developed using VP-EXPERT, a rule-based expert system development tool. Because of memory space limit, it is actually composed of two knowledge bases, one for determining APRM operating status and the other for checking Tech Spec 3.1.C compliance. Totally, there are 60 rules in which 42 are related to the Oyster Creek Technical Specification stated above. At the beginning of the consultation, the user is asked about the status of the APRMs and the LPRM detectors. A menu of APRM channels and LPRM locations is presented for the user to select the bypassed or failed detectors. Once the detector configuration has been entered, the expert system would determine the APRM channel status and check whether Tech Spec 3.1.1.A and 3.1.B are complied. Message will be printed for noncompliance situation. The user is then asked whether to continue for Tech Spec 3.1.C compliance check. Sample detector configuration and the corresponding ESAO output are given in Figures 9 and 10.

CONCLUSION

These two expert systems will be used during cycle 12 start-up which is scheduled for Spring, 1989. It is expected that the usage will demonstrate that expert systems can be used to support plant operation. Prior to Oyster Creek Cycle 12 start-up, SMARTRODS was used to generate the control rod withdrawal sequence table. The form input was found to be very easy to us. After a demonstration session, the core engineers were able to use it without any difficulty. Because of the change in operation strategy which is not reflected in the move rules, minor adjustments of the sequence table were required. This was done manually by the reactor engineer, with the revised sequence table checked by the expert system. The running time for SMARTRODS is about five minutes depending on the control rod patterns. Using SMARTRODS, a control rod withdrawal sequence table can be generated and checked in 10 minutes. This saves two to three days of a reactor engineer's time if the table has to be generated and checked manually. The capability of providing a quick and thorough check of the revised sequence table during start-up will be very useful. Using ESAO, the operator can check technical specification compliance for alternative detector configurations when it is necessary to bypass an APRM channel or LPRM detectors. The running time for a consultation session is about three minutes regardless of the detector configuration. Compared with the time needed for manual determination, i.e. two to five minutes for simple cases and half to an hour for complicated cases, the use of ESAO could be a very useful tool for the reactor operators and the reactor engineers during the start-up. In summary, the expert systems will facilitate the decision making during start-up. The actual benefits will be evaluated during Cycle 12 start-up.

Both SMARTRODS and ESAO can be written using conventional programming style. We chose the expert system approach because it gives clearer knowledge representation and is easy to modify. In addition, we would like to

APRM channel bypassed: 1



X: failed or inoperable detector

Figure 9

Sample Detector Configuration

```
      AN EXPERT SYSTEM TO DETERMINE OC PROTECTION SYSTEM STATUS
            BASED ON THE FAILED DETECTORS AND BYPASSED APRM

********************     USER      INPUT    ************************
APRM channel 1 is bypassed in the RPS-1
There is NO APRM channel bypassed in the RPS-2
Level A detector failed in the following LPRM :
 LPRM : X36Y17
 LPRM : X20Y25
 LPRM : X12Y17
 LPRM : X20Y33
Level B detector failed in the following LPRM :
 LPRM : X28Y25
 LPRM : X28Y33
Level C detector failed in the following LPRM :
 LPRM : X36Y17
Level D detector failed in the following LPRM :
 LPRM : X44Y25
 LPRM : X12Y17
********************************************************************

Channel 4 is FAILED because both level A & level C of
the same LPRM are failed (Tech Spec 3.1.B.2)

********************     APRM      STATUS    ***********************
*                                                                  *
                          RPS - 1                  RPS - 2
      Quadrant      Channel    Status         Channel    Status

        1st            1       BYPASSED          5       OPERABLE
        2nd            2       OPERABLE          6       OPERABLE
        3rd            3       OPERABLE          7       OPERABLE
        4th            4       FAILED            8       OPERABLE
*                                                                  *
********************************************************************


********************    TECH SPEC STATUS    ************************
More than 1 APRM channel failed in RPS-1 or RPS-2.
Not in compliance with Tech Spec 3.1.A.2 & Table 3.1.1
HIT /Q to end the consultaion or Y for Tech Spec 3.1.C.1 compliance.


Tech Spec 3.1.C is not satisfied because LPRM's located at X28Y33 and
X28Y25 have more than 3 detectors failed in either A&B or C&D levels.
End of consultation.
```

Figure 10
Sample ESAO Output

investigate the potential usage of expert system to support plant operation. Our experience shows that for an expert system to be accepted as a useful tool, it must have a good user interface, allowing the user to start consultation without any specific training. Otherwise, it will be very difficult to attract the user to overcome the initial learning stage. Another desirable feature is to print the input and output data in the same format as used in plant operation procedure, thus reducing the paper work. This is an area of future improvement for SMARTRODS and ESAO. We also plan to modify the move rules in SMARTRODS to reflect the change of Oyster Creek operation strategy. The company currently has no plan to develop large expert systems, but we will continue our efforts in developing small expert systems, using available development tools to support plant operation.

ACKNOWLEDGMENTS

I would like to acknowledge J. Reierson of MITRE Corporation for the initial work on RODS which introduced expert system for plant operation support at Oyster Creek. I am very grateful to D. Cain of EPRI for his assistance in developing the RODS user interface. Both expert system would not have been developed without the active cooperation of R. V. Furia of GPU Nuclear to serve as the domain expert. I gratefully acknowledge his contribution. I also want to thank D. Notigan, J. Sedar, H. Sharma, and R. Thompson of Oyster Creek Core Engineering for testing the expert systems.

REFERENCES

1. J. Reierson, An Application of Expert System Technology to Nuclear Power Plant Operations, McLean, Virginia: The MITRE Corporation, MTR-83W196, October 1986.

2. D. Cain, Small Artificial Reasoning Tookkit, Palo Alto, CA: Electric Power Research Institute, EPRI NP-5646-CCML-SR, 1988.

3. Oyster Creek Nuclear Generating Station Operating License and Technical Specifications, GPU Nuclear and Jersey Central Power & Light Company.

4. VP-Expert Rule-Based Expert System Development Tool, Paperback Software, 1987.

Residual Heat Removal System Diagnostic Advisor

LLOYD TRIPP
Artificial Intelligence and Sensor Engineering
1515 South Manchester Avenue
Anaheim, California 92802-2907, USA

ABSTRACT

The Residual Heat Removal System (RHRS) Diagnostic Advisor is an expert system designed to alert the operators to abnormal conditions that exist in the RHRS and offer advice about the cause of the abnormal conditions. The Advisor uses a combination of rule-based and model-based diagnostic techniques to perform its functions. This diagnostic approach leads to a deeper understanding of the RHRS by the Advisor and consequently makes it more robust to unexpected conditions.

The main window of the interactive graphic display is a schematic diagram of the RHRS piping system. When a conclusion about a failed component can be reached, the operator can bring up windows that describe the failure mode of the component and a brief explanation about how the Advisor arrived at its conclusion.

The RHRS Diagnostic Advisor was developed using the Automated Reasoning Tool (ART) from Inference Corporation running on a Symbolics 3675.

INTRODUCTION

The Residual Heat Removal System (RHRS) Diagnostic Advisor is an expert system developed under contract to the Department of Energy and in conjunction with Impell Corporation and the Commonwealth Edison Company. The RHRS Diagnostic Advisor is intended to demonstrate how expert systems technology can be used to support some aspects of RHRS operation particularly system monitoring and off-normal condition diagnosis. While the RHRS Advisor is being developed for the nuclear industry in general, it is modeled after the RHRS at the Zion nuclear power plant operated by Commonwealth Edison Company. Where possible, the information given here is generically applicable to Westinghouse-designed RHR systems. However, in order to make the Advisor functional for the Zion plant, the majority of this data is Zion-specific. Before the details of the Diagnostic Advisor are presented, a brief description of the RHRS and its operation will help the reader appreciate why the RHRS was chosen for this expert system technology demonstration.

The RHRS is a major component of the decay heat removal system in a nuclear power plant. Even after the nuclear chain reaction is stopped, there is a significant amount of heat produced by the continuing radioactive decay of the fission products. The decay heat removal system, as the name implies, is designed to remove this remaining decay heat. When the Reactor Coolant System (RCS) conditions approach 350 F and 425 psig, the RHRS is connected to the RCS to continue the heat removal process until cold shutdown conditions are reached. Once in cold shutdown, the RHRS continues to transfer heat to the Component Cooling Water (CCW) system to maintain stable cold shutdown conditions. Conversely, the RHRS can also be aligned to permit heatup of the RCS from cold shutdown conditions in preparation for plant startup.

In the Zion nuclear power plant, the RHRS is required to perform several other functions as well depending on the mode of plant operation. In the event of a loss of coolant accident, it provides low pressure injection of borated water into the RCS cold legs and can subsequently be realigned to recirculate reactor coolant and provide containment spray from the containment recirculation sump. The RHRS is also employed to transfer refueling water between the Refueling Water Storage Tank (RWST) and the refueling cavity before and after refueling operations.

Although decay heat removal at first glance appears to be a relatively benign power plant function, it has recently come under a great deal of scrutiny. For example, the Nuclear Regulatory Agency (NRC) has identified shutdown decay heat removal as an Unresolved Safety Issue (A-45). The Nuclear Safety Analysis Center (NSAC) has published two reports summarizing their safety analysis of the RHRS for pressurized water reactors (1) and boiling water reactors (2). The NSAC reports state:

> Reduced decay heat levels present during these [safety] events usually permit more time to respond to problems than is

available during power operation. However, since fewer automatic protective features are operative during cold shutdown, both prevention and termination of these events depend heavily on operator action.

Residual Heat Removal System Diagnostic Advisor is designed to provide information and advice to the operators so they can perform the proper action. The Advisor's role, will be that of a tireless "noticer" of discrepancies, and a judicious "presenter" of possible diagnoses. It will not attempt to override the operator's judgement. Rather, it will make its own reasoning process transparent enough to the operator so that potential violations of common sense can be detected and overridden by the operator. In this way, the Advisor will make a positive contribution to the operator's capacity, without disabling the component of human reason and common sense so essential to plant control and safety.

With this description in mind, some boundary must be placed on the detail of the knowledge that is to be encoded in the expert system, and on the scope of the off-normal conditions that it should be able to correctly diagnose.

SCOPE OF THE RHRS DIAGNOSTIC ADVISOR

The scope of the RHRS Diagnostic Advisor can partially be defined in terms of the breadth and depth of the off-normal conditions that it should be able to correctly diagnose. The breadth means the number and type of off-normal conditions, while the depth means the level of detail to which it can analyze and explain the off-normal conditions. The current design of the Advisor is intended to provide a satisfactory compromise between the breadth and depth.

In terms of the breadth, the Advisor is designed to recognize single-point failures of the flow-control components as well as abnormal sensor behavior. By flow-control component, I mean all the valves in the RHRS and the two RHRS pumps. The sensors include all the flow, pressure, temperature, and level sensors that are part of the RHRS. Note that the breadth specifically excludes pipe failures inside the RHRS and component failure outside the RHRS that adversely affect RHRS operation (except for a limited number of specific cases). Even though the RHRS is often thought of as an isolated system, it is in fact coupled with all the other systems that comprise the nuclear power plant. This coupling with the other systems makes the definition of the breadth somewhat arbitrary. It does, however, result in a breadth that covers a large number of off-normal conditions yet is still of manageable size so that sufficient depth can be included in the scope.

The depth of the scope is limited to the identification of the component causing the off-normal conditions and the reasons the Advisor believes the component is causing the off-normal conditions. If the reasoning process does not result in the identification of a single component, then members of the final set of suspected components are identified. This depth specifically excludes identification of subcomponents. This means, for example, if a motor-operated valve is malfunctioning, the Advisor is not designed to determine if it is due to shaft seizure or actuator motor failure.

The scope also includes recognition of the wide range of conditions that are considered normal operation. Without including this in the scope, it would be very difficult to distinguish between normal and off-normal conditions.

The RHRS Diagnostic Advisor is not designed to directly manipulate system components, such as motor-driven valves, either to test its failure hypotheses or to implement repair actions. This reflects our philosophy that a human being should be "in the loop" at all times, with the system merely adding its perceptions to the operator's and giving the operator advice.

In part because of its several operating alignments, some off-normal conditions in the RHRS are unobservable until an alignment change makes them observable. For example, if a manually-operated valve, which has no position sensors, is in an incorrect position, then its off-normal condition will remain unobservable until the RHRS is aligned in such a way that the normal flow of coolant is changed by the mispositioned valve. A condition can also be unobservable due to limitations of the RHRS sensors and the frequency at which the sensor measurements are sampled. An example of a sensor limitation is that there does not exist a direct measurement of the position for the air-operated butterfly valves, only a demanded position. The frequency at which the sensor measurements are sampled sets an upper limit on the observability of some oscillating conditions. Currently, the Zion plant computer samples sensor readings about once a minute. Suffice to say that the Advisor will only be able to diagnose disorders that are observable.

RHRS DIAGNOSTIC ADVISOR SYSTEM ARCHITECTURE

There are several knowledge-based techniques for performing problem diagnosis. Each of these techniques tries to encode the knowledge an "expert" uses to diagnose problems in some form, and to apply this encoded knowledge to the set of problems covered by the knowledge. The knowledge contained in the expert system and how it is encoded determines, to a large extent, the ability of the Advisor to diagnose off-normal conditions within its scope. There were several primary sources of knowledge used to develop the knowledge base for the Advisor. The experts at Impell Corporation provided the following printed information:

- a description of the RHRS, its components, and a schematic diagram,
- a description of recent safety events in the nuclear power industry involving the RHRS,
- an extensive table of component failures and their associated sensor indications, symptoms, and proper operator response,
- a summary of pertinent Technical Specification limits and Zion Station procedural precautions, and
- a summary of the normal operating procedures for the Zion Station RHRS.

Experts from Impell were also used throughout the project as a source for answers to technical questions about the RHRS and the use of expert systems in the control room.

Personal interviews were conducted with control room engineers and operators from Commonwealth Edison to get a first hand account of the diagnostic support that could be used in the control room. Concepts for the user interface were also discussed.

The experts at Impell ran 15 test cases on the power plant simulator at Zion Station, to gather simulated sensor data from the RHRS so that it could be used test and partially validate the Advisor.

Using these sources of information, it became clear that a great deal of knowledge about the physics of the piping system and the causal relationships of one action to another are needed in order to detect and diagnose the off-normal conditions that can be present within the scope of the RHRS Diagnostic Advisor. For this reason, a architecture combining model-based and rule-based reasoning is used. Each of these reasoning techniques has both strengths and limitations when used in a diagnostic expert system. Combining the two techniques can lead to better system performance.

Model-based Reasoning

One technique that can be used to encode the physics of the RHRS piping system and the causal relationships between one action and another is called model-based reasoning. The idea behind it is similar to building mathematical models to describe physical systems except that rather than formulating a precise QUANTITATIVE model, a less precise, more intuitive QUALITATIVE model is used. Just like the mathematical model (a set of differential equations), the level of abstraction used by the qualitative model depends on how the model is going to be used or what it is trying to predict. For example, when analyzing an electric circuit, a common level of abstraction is to model the resistors, capacitors, and inductors as PURE resistors, capacitors, and inductors even though the actual physical components have varying amounts of all of these properties. Likewise, if we are only interested in determining if the flow through a segment of pipe is adequate or not, a detailed model of the cross-sectional velocity flow profile is not needed. This is because we can assume that the RHRS was designed so if all the components are functioning properly and are properly aligned, there will be adequate flow. The level of abstraction used in the Advisor's qualitative model, then, is such that it can reason about whether the components are functioning properly and are properly aligned.

Another modeling abstraction that is commonly used for systems of connected components, is to model the behavior of the entire system as the aggregate of the behaviors of the individual components that comprise the system. The reason for this abstraction is that modeling the behavior of a complex system as a whole is much more difficult than modeling the behavior of its components and linking them together. The system behavioral model resulting from linking the behaviors of its components will not be exactly the same as a model of the system as a whole (due to interactions of components that are not accounted for when the component behaviors are linked together) but it should be accurate enough to detect the types of off-normal conditions defined in the scope of the Advisor. This modeling abstraction will be used here when the qualitative behavior of sections of the RHRS is determined by the aggregate of the qualitative behaviors of the individual components that comprise the section. For instance, the behavior of the components that comprise the A train of the RHRS determines the behavior of the A train (as long as the A and B trains are isolated). If Pump A stops pumping, it determines that there will be no flow down the A train. So the individual component, Pump A, can determine the behavior of a section of the RHRS, the A train.

Because qualitative models are simplified to the point of being almost intuitive, the reasoning process that uses these models more closely follows the human reasoning process. Qualitative models make more use of symbols and relative values rather than numbers. This is because humans can better handle symbols rather than the numbers from a quantitative or numerical model.

Model-based reasoning also makes the causal relations in the system more explicit to the human than a set of equations. People often use causal relations to diagnose problems. If an automated reasoning system like the RHRS Diagnostic Advisor is to diagnose problems and explain its reasoning process to people, then that reasoning process should be close to what people use or the reasons will not make sense. Causal relations connected by the flow of coolant through the piping will be used extensively since this is the major causal link between actions that take place in different parts of the system.

The robustness of the representation is another strong point for using model-based reasoning to encode the knowledge needed to solve the problems of the RHRS. Because the models are qualitative representations of the components and the causal relations between them, they have a better foundation in the physics of the system than an encoding scheme that does not make this link explicit. This foundation in physics gives the Advisor a deeper understanding about the RHRS which improves its monitoring and diagnostic tasks.

A more in-depth treatment of model-based reasoning and qualitative physics can be found in (3).

While model-based reasoning alone may initially seem adequate for all aspects of the RHRS Diagnostic Advisor, it does have some limitations. One limitation is that model-based reasoning is so well suited to reasoning about causal relationships between facts that it is not well suited to reason when no causal relationship exists. Another limitation is that useful heuristics or "rules-of-thumb" do not fit well into the model-based reasoning scheme. Fortunately, these limitations are the hallmark of rule-based reasoning.

Rule-based Reasoning

Rule-based reasoning is the technique most often associated with expert systems. This technique is the foundation of classic expert systems such as Mycin and Xcon.

Rule-based reasoning, however, is not suitable for the diagnostic tasks of the Advisor. This is because the rules are not based on the physical structure of the RHRS. The result is the rules have no ability to reason beyond the specific symptom-fault cases that are explicitly defined. If, due to some oversight, the rule covering a symptom-fault case was left out, a rule-based system would not provide the correct diagnosis. Also, a slight variation in the symptoms for a known fault may preclude the intended rule from firing so that no diagnosis could be made. This is referred to as "falling off the knowledge cliff."

While not well suited to the diagnostic tasks of the Advisor, rule-based reasoning is well suited to perform other important tasks such as:

- mapping the numerical sensor readings to the symbolic values used by the model-based reasoning system,

- monitoring the sequence of events and operator actions performed while changing the valve alignment, and

- handling the intelligent man-machine interface.

A good discussion of the trade-offs between model-based and rule-based diagnostic techniques was presented at (4).

With this architecture in mind, a functional description of the advisor will illustrate how the reasoning techniques are utilized to detect and diagnose off-normal conditions in the RHRS.

FUNCTIONAL DESCRIPTION OF THE RHRS DIAGNOSTIC ADVISOR

The RHRS Diagnostic Advisor has two main functions:

1. to monitor the data coming from the sensors and from the operator to determine if something is wrong, and

2. if something is wrong, to determine the cause of the situation and explain it to the operator upon request.

Both of these functions are implemented using the model-based reasoning technique as its basis.

Monitoring

For most of the time, the Advisor will be silently performing its monitoring function looking for indications that the RHRS is not functioning correctly. The technique used for detecting off-normal behavior is based on the concept of "expected state violations." The concept is that each component needs to be in its expected state if the system is going to be declared operating normally. If a component is not in its expected state, i.e. its expected state is violated, then the off-normal behavior has been detected.

The state of a component describes the operating condition of the component to the level of abstraction used by our qualitative models. For motor-operated valves, the states include {OPEN, CLOSED, INDETERMINANT}. For the pumps, the possible states are {ON, OFF}. The state of most sensors will be one of {LOW, NORMAL, HIGH}. The process of mapping switch readings and sensor readings to states with absolute qualitative values such as OPEN, CLOSED, ON, and OFF is trivial. However, the process of mapping numerical sensor readings to states with relative values such as LOW, NORMAL, and HIGH is much more difficult. The Advisor uses a dedicated set of rules for each sensor to perform this mapping. The mapping rules use the current value of the sensor as well as information about the current alignment, trend (rising, falling, or steady), and changes in the state of other components that can affect the sensed value.

The expected state of each component is stored in a record-like structure (called a schema in the ART language) for the current valve alignment. For most components, the expected state is a single value that is determined in advance. For example, we know from the alignment procedures which motor-operated valves are expected to be CLOSED and which ones are expected to be OPEN. Some expected states cannot be determined with certainty in advance because the operator has some discretion as to what the expected state will be. For example, in the Cooldown alignment, the operator determines which of the two pumps to start or whether to start both of them.

When a new data item comes in from a switch, sensor, or other source, rules will fire which take the data item and compare it to its currently expected state. If the values are the same, then the monitoring function continues to check other data items that may have come in. If the values conflict, then the operator is notified that an expected state violation exists that will be further examined by the diagnostic rules.

Diagnosis

The diagnostic rules establish a link between the expected state violation and the knowledge about the structure and the causal relationships present in the RHRS. The use of causal relationships is particularly useful when trying to resolve expected state violations of components that affect the coolant flow through the system. Since this involves most of the components, we can expect that an examination of the causal relationships linked by flow will greatly aid in determining which component is violating its expected state and HOW it is violating its expected state.

The diagnosis proceeds by using the causal relationships encoded into the Advisor's data structures, models, and functions to find a set of components that could possibly be causing the unexpected component state. A separate set of suspected components is generated for each component that is violating its expected state. Once all the sets are complete, the sets are intersected to try to find common components to all the sets. If the set resulting from the intersection still contains more than one component, then other rules are used to gather redundant information on the state of the components to aid in further reducing the number of suspected components.

The use of redundant information present in the RHRS makes this technique robust. For example, the state of a motor-operated valve can be ascertained by its position limit switches as well as by determining if there is flow on either side of it. Likewise, the flow sensors provide redundant information about the flow in the RHRS during many conditions.

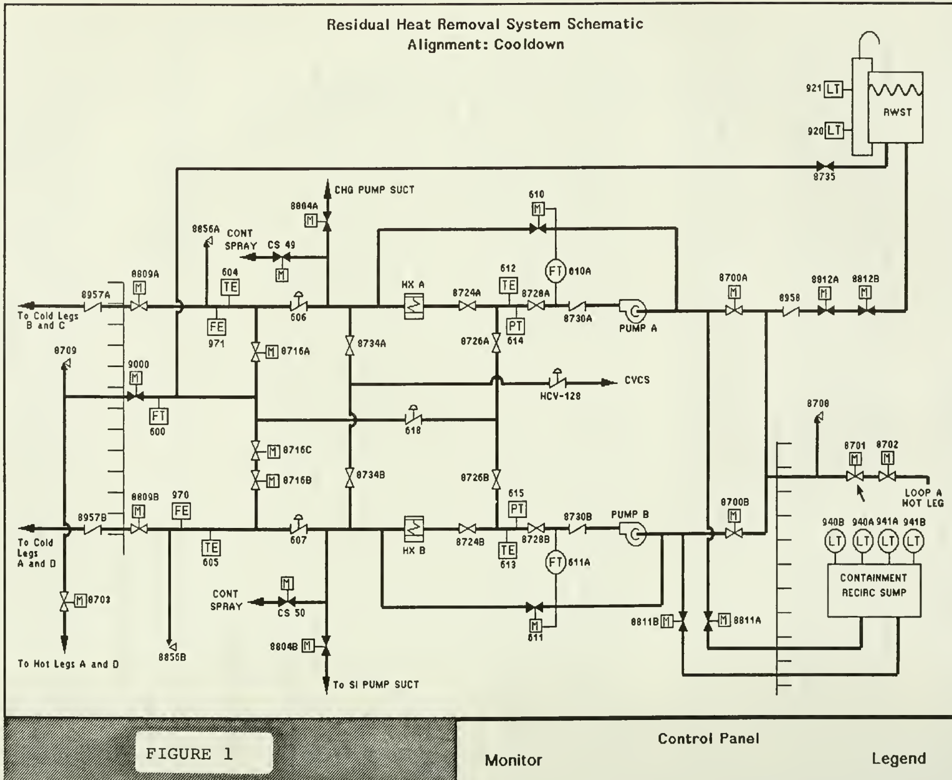
Faults associated with pressure and temperature have similar causal relationships that can help in identifying the component responsible for the expected state violations.

Sometimes it is not possible to identify a single component that is responsible for the observed expected state violations. In this case, the Advisor identifies a ambiguity group to the operator and asks the operator questions that could help to resolve the ambiguity. The answer to the questions may involve the gathering of additional information through local inspection. Ambiguity groups can arise due to the limited observability of the system given the sensors that are present. Potentially large ambiguity groups can arise if data from a sensor becomes unavailable (for instance, due to repair). In this case, the Advisor will rely even more on the operator to answer questions that can reduce the size of the ambiguity group. Once a component (or an ambiguity group) has been identified as the cause of the expected state violations, the Advisor will explain a summary of its reasoning process to the operator so that he can use this information as a "common sense" check of the result.

OPERATOR INTERFACE

The operator interface is based on the schematic diagram of the RHRS (Figure 1). It serves as the focal point for all interaction between the operator and the RHRS Diagnostic Advisor.

The schematic diagram is not a static presentation like a schematic drawn on paper. Rather, it is updated with information to show the current valve alignment and uses animation to show which pipes have coolant flowing through them.


Residual Heat Removal System Schematic
Alignment: Cooldown
RWST
921 LT
920 LT
8735
CHG PUMP SUCT
8804A
8856A
CONT SPRAY
CS 49
610
8809A
8957A
To Cold Legs B and C
604 TE
FE 971
606
HX A
8724A
612 TE
8728A
FT 610A
8730A
PUMP A
8700A
8958
8812A
8812B
8716A
8734A
8726A
PT 614
CVCS
HCV-128
8709
9000
FT 600
618
8708
8716C
8716B
8734B
8726B
8701
8702
LOOP A HOT LEG
8809B
970 FE
8957B
To Cold Legs A and D
TE 605
607
HX B
8724B
615 PT
TE 613
8728B
FT 611A
8730B
PUMP B
8700B
940B 940A 941A 941B
LT
CONTAINMENT RECIRC SUMP
8703
CONT SPRAY
CS 50
611
8811B
8811A
To Hot Legs A and D
8856B
8804B
To SI PUMP SUCT
FIGURE 1
Control Panel
Monitor
Legend

The Advisor emphasizes the interactive exchange of information between the operator and the Advisor by providing mechanisms so that the operator can request information as well as respond to questions asked by the Advisor. This differs considerably from the sources of information that are currently available to the operators. Most of this information comes from the meters and status lights that are mounted on the control board. Supplementary displays using CRTs present a small number of reactor parameters that the user can select for display. None of these devices ever ask for information from the operator; they are output only.

The operator will interact with the Advisor exclusively through the use of the mouse pointing device. This means when the operator wants to request information about a component, he points and clicks the mouse on the component. Figure 2 shows the operator display after the operator has requested the time history of flow element 971, temperature element 604, and pressure transmitter 614. The data shown in the strip chart displays is the actual data from the early phase of one of the component failure scenarios simulated on the Zion Station control room simulator. This particular failure scenario involves one of the RCS pressure transmitters (PT-405) failing HIGH approximately 20 minutes into the scenario. Due to a safety interlock, this pressure transmitter failure causes the hot leg suction valve 8702 to close. The closing of valve 8702 causes the coolant flow in the RHRS to stop. This, in turn, causes the expected state of 8702, the flow, pressure, and temperature sensors to be violated. The diagnostic rules and functions use the causal information to determine that the root cause of the off-normal conditions is the failure of PT-405. Figure 3 shows the display after the failure has occurred and after the operator has moused on the ATTENTION icon. By mousing on the ATTENTION icon, the operator gets a terse textual message in a "pop-up" window describing the reason it highlighted the component. Note that because PT-405 is not a part of the schematic display of the RHRS, the Advisor highlights the area around the text "LOOP A HOT LEG" to indicate that the suspected component is a part of the RCS.

The strip chart displays can be brought up for each of the sensors shown on the schematic diagram. The operator can configure the strip charts in many ways to show the information he wants in a form that is easy to interpret. The strip charts can be configured in the following ways:

- can be hidden or exposed by mousing on the sensor icons,
- the vertical axis can be rescaled and the low offset from 0,

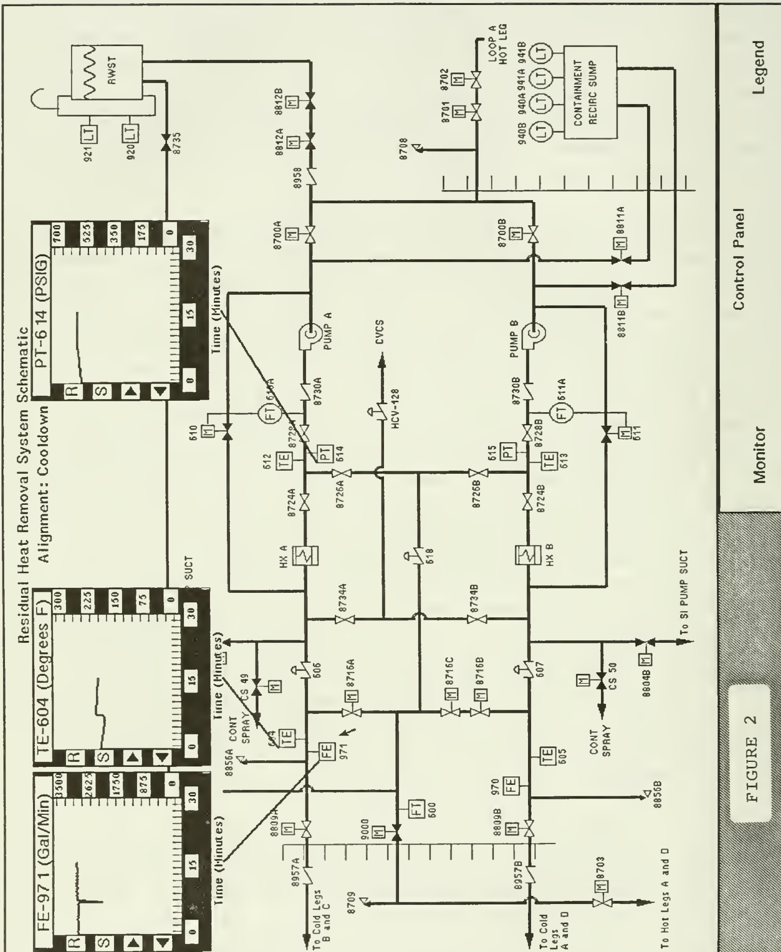
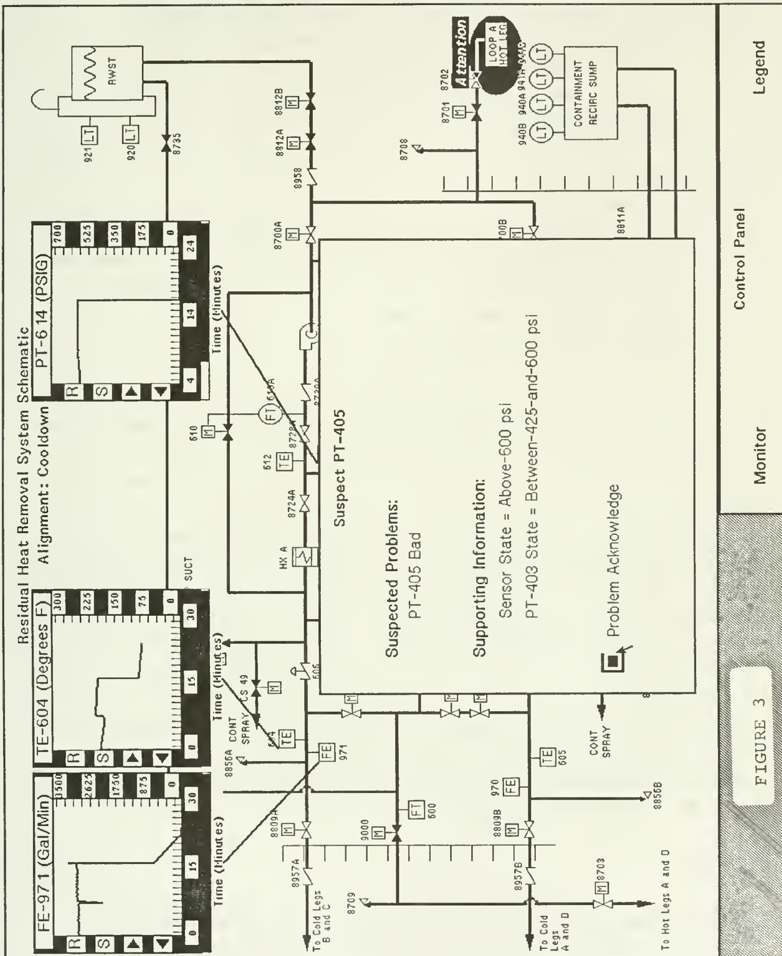

Residual Heat Removal System Schematic
Alignment: Cooldown
FE-971 (Gal/Min)
TE-604 (Degrees F)
PT-614 (PSIG)
Time (Minutes)
RWST
CONTAINMENT RECIRC SUMP
PUMP A
PUMP B
HX A
HX B
CVCS
To Cold Legs B and C
To Cold Legs A and D
To Hot Legs A and D
To SI PUMP SUCT
LOOP A HOT LEG
CONT SPRAY
Monitor
Control Panel
Legend

FIGURE 2



Residual Heat Removal System Schematic
Alignment: Cooldown
FE-971 (Gal/Min)
TE-604 (Degrees F)
PT-614 (PSIG)
Time (Minutes)
RWST
Suspect PT-405
Suspected Problems:
PT-405 Bad
Supporting Information:
Sensor State = Above-600 psi
PT-403 State = Between-425-and-600 psi
Problem Acknowledge
Attention
LOOP A HOT LEG
CONTAINMENT RECIRC SUMP
CONT SPRAY
To Cold Legs B and C
To Cold Legs A and D
To Hot Legs A and D
Monitor
Control Panel
Legend

FIGURE 3

- the horizontal time axis can be rescaled to show more data points or rescaled to zoom in on a time segment of interest,
- the charts hold eight hours of data so the operator can scroll back and forth in time, and
- the charts can be moved to any location on the schematic and even overlap each other.

Host Environment

The RHRS Diagnostic Advisor is currently hosted on a Symbolics 3675 computer. The Advisor is implemented using the Automated Reasoning Tool (ART) expert system shell supplemented by CommonLISP functions. The Symbolics has a special hardware architecture for performing symbolic computation. This makes it an ideal host for developing and testing the Advisor.

Before sensor data is sent to the Advisor, it needs to be preprocessed to put it into the form of a list with a descriptive label. This way the Advisor will have no problems determining what sensor the data came from. The Symbolics computer receives its data from the sensor preprocessor via the Symbolics Ethernet port. The computer performing the sensor preprocessing is a Sun 3/160. The Sun is a fast, general purpose workstation that can easily perform the task of preprocessing the sensor data used to test the Advisor. The Ethernet link between the Sun and Symbolics was already used for the purpose of sending data processed on the Sun to the Symbolics so little extra development work is required to use the link for this purpose.

The ART expert system shell is used in the development and testing of knowledge-based systems like the RHRS Diagnostic Advisor. Built into the shell are the necessary tools for developing the data structures and rules that hold the knowledge about the RHRS. It also has a graphic interface tool for creating the graphic-based operator interface.

FUTURE WORK

The RHRS Diagnostic Advisor is a prototype which must under go an extensive amount of testing, verification, and refinement before it can be used in a control room. The control room simulator is an ideal place to continue the development of the Advisor because many failure scenarios that simulated to test the Advisor. Another advantage of using the control room simulator is that the operators that are using the simulator for training can be exposed to the Advisor in an environment where they would be willing to experiment and use the Advisor. Valuable feedback on the man-machine interface could also be gained.

The version of the ART expert system shell used to develop the Advisor is probably not suitable for use in an attached diagnostic system that must run continuously for long periods of time. Also, there is no easy way to strip away the software development tools to get a small executable image and prevent the operators from modifying the Advisor software. The C language-based expert system shell called ART-IM will be evaluated to see if it is better suited to the attached system environment.

Since the configuration of the RHRS is similar to other nuclear plant piping systems we are anticipating the development of other expert systems as advisors for systems such as the Emergency Core Cooling System, Feedwater System, Component Cooling Water System, and Service Water System.

CONCLUSION

The RHRS Diagnostic Advisor has demonstrated that expert systems can be used to support some aspects of RHRS operation by having on-line expert advice. The Advisor also demonstrated the performance of using a combination of model-based and rule-based techniques for diagnosing problems with piping systems like the RHRS. The advanced man-machine interface demonstrates how large amounts of information can be made available to the operators without overwhelming them.

ACKNOWLEDGEMENTS

I would like to acknowledge the support of Victoria Walther and Lori Armstrong of Impell Corporation for providing the expert information used in this project.

REFERENCES

1. Residual Heat Removal Experience Review and Safety Analysis, Pressurized Water Reactors, NSAC-52, Palo Alto, CA: Nuclear Safety Analysis Center, January 1983.

2. Residual Heat Removal Experience Review and Safety Analysis, Boiling Water Reactors, NSAC-88, Palo Alto, CA: Nuclear Safety Analysis Center, March 1986.

3. Bobrow, Daniel G., ed. Qualitative Reasoning About Physical Systems. Cambridge, MA: The MIT Press, 1985.

4. Diagnostic Expert Systems in Aerospace and Defense Electronics, Los Angeles, CA: Institute for International Research, June 27-28, 1988.

A PC-based Expert System for Nondestructive Testing

R. SHANKAR, R. WILLIAMS, C. SMITH, and G. SELBY
EPRI NDE Center
Operated by
J. A. Jones Applied Research Co.
1300 W. T. Harris Blvd
Charlotte, North Carolina 28213 USA

ABSTRACT

Rule-based decision logic which can emulate problem-solving expertise of humans is being explored for power plant nondestructive evaluation (NDE) applications. This paper describes an effort underway at the EPRI NDE Center to assist in the interpretation of NDE data acquired by automatic systems during ultrasonic weld examination of boiling-water reactors (BWRs). A personal computer (PC)-based expert system "shell" was used to encode rules and assemble knowledge to address the discrimination of intergranular stress corrosion cracking (IGSCC) from benign reflectors in the inspection of pipe-to-component welds. The rules attempt to factor in plant inspection history, ultrasonic examination data and, if available, radiography testing data; a majority of them deal with specific ultrasonic signal temporal and spatial behavior during automatic scanning. The difficulties in interpretation are due to the similar ultrasonic signal response from IGSCC and weld geometrical reflectors, such as roots and machined counterbores.

The expert system is configured in a question-answer format and consists of approximately 300 decision rules.

The expert system has been integrated on a PC with a "feature-based" imaging system capable of acquiring, displaying and computing image features pertinent to the consultation. The integrated capability was achieved using commercially available and EPRI-developed products. The system was evaluated at the EPRI NDE Center on field-removed samples with service-induced IGSCC and is currently being evaluated by utilities.

The paper describes the efforts in the development of the expert system.

OVERVIEW

IGSCC of piping in boiling-water reactors (BWRs) first received attention in the U.S. in 1975 when all the BWRs were shut down for inspection of welds in several piping systems. Later in 1982 IGSCC was discovered in larger diameter pipes (1). Numerous ultrasonic "indications" were observed in the inside surface region near the welded area, and industry took steps to deal with the problems. These steps included augmentation of existing inspection guidelines, more detailed inspection procedures and control of water chemistry to inhibit initiation of IGSCC.

The EPRI Nuclear Power Division initiated an effort at the EPRI NDE Center in 1988 to capture and codify expert knowledge used in the interpretation of ultrasonic testing (UT) data during BWR weld examination. Difficulties in data interpretation arise because of the close resemblance of the signatures from cracks and other geometrical reflectors in the weld region. While proper instrumentation and careful adherence to experimental procedures play a large role, experiential knowledge of the problem was determined essential for data interpretation. Earlier attempts to implement a "purely algorithmic" approach yielded mixed results; they were sometimes too rigid to perform satisfactorily on samples outside the training set. It was long recognized that operators considered past weld history as well as evidence from other, auxiliary NDE techniques -- such as radiographic testing (RT) -- to arrive at an overall decision. A first attempt was made in 1986 to identify common rules used by operators in ultrasonic data interpretation. These rules and pictorial illustrations were published in an EPRI report in 1988 (2).

Recent advances in computer hardware and software and the proliferation of low-cost expert system "shell" programs made it possible to consider such systems for symbolic and numerical data manipulation. Rules were developed initially to interpret ultrasonic B- and C-scan image data with the information documented in (2). It was assumed that the operator could view these images during consultation. The questions related to UT and RT data required the user to accurately assess the inspection data. The questions were restricted to a qualitative appraisal of the relevant UT image data: was the UT indication length "short" or "long"? Are the reflector echodynamics "narrow" or "wide"?

The evaluation of the first prototype was conducted by one of the authors on field-removed pipe specimens with service-induced IGSCC and field-quality geometrical reflectors and was satisfactory. However, in another independent evaluation by an NDE Center staff member, the system performance was considerably worse. The difference in performance was attributed to the difference in familiarity with questions and questioning style. Specifically, it was concluded that improvements were needed in:

- the clarity and completeness of the questions and instructions.

- the graphics used to aid in answer selection, especially for those questions that required a qualitative answer (how narrow is "narrow"? for example); and

- the inclusion of questions asked on weld history and the weighting assigned to the RT data.

These recommendations led to a major revision in early 1989 wherein some rules and questions were modified and weld history rules were added to provide information on the historical evidence.

Figure 1 shows an overview of the BWR weld examination expert system. The consultation is conducted in three major areas: weld history, UT data and RT data. The system responds with evidence of cracking based on weld history and on NDE data. The historical and NDE data evidence are not combined (See Figure 1). Future revisions will consider rules to combine historical and NDE data evidence. Six questions are asked pertaining to weld history. These questions relate to cracking in sister units and in other components; prior inspection findings on the component; stainless steel material type and component configuration. The questions on UT and RT data consider detailed characteristics and assume ability to view the UT image data. This capability was provided wherein the user could operate under a "windows" environment and toggle among the consulting sessions, a UT imaging and analysis program that could display and compute mathematical "features" pertinent to the consultation, and an ultrasonic ray tracing package that allows the user to postulate different inspection scenarios for the component under inspection.

The product will continue to be evaluated by the NDE Center as well as by three utilities and a vendor. The main purpose of this evaluation is to determine system functionality, accuracy of questions asked, and the need for additional questions and rules to combine knowledge. The purpose is <u>not</u> to demonstrate system performance. The expected results from this evaluation will include improvements in man/machine interface and incorporation of additional rules and plans for future deployment.

BWR WELD INSPECTION

Ultrasonic inspection of these welds is performed either manually or automatically and is conducted during a plant outage. In manual inspection, the operator "scrubs" the pipe with a contact transducer, usually operating in pulse-echo mode, and observes the response on a calibrated display. In automatic inspection, a

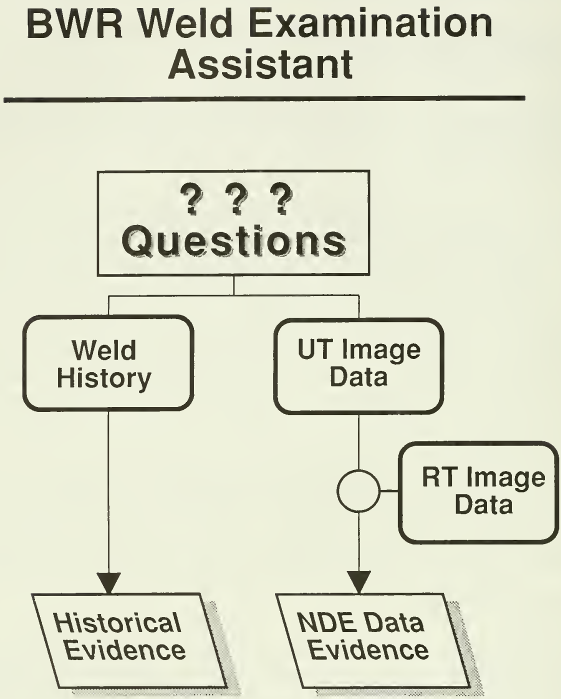


Figure 1. Overview of BWR Weld Examination Expert System.

transducer manipulator scans the pipe according to programmed instructions as ultrasonic data are acquired and stored during the scan pattern. The data are subsequently imaged and analyzed. Automatic inspection is preferred because modern computing platforms are powerful and economical, and weld data can be well documented and compared between plant outages. In addition, with more emphasis placed on reducing total plant radiation exposure, automatic systems are preferred over manual methods. Manual inspection is performed when weld accessibility is limited and to confirm automatic inspection results.

The cracking occurs on the inside surface, close to the weld in the heat-affected-zone (HAZ). Difficulties in detection of IGSCC by ultrasonic means are primarily due to the close resemblance of IGSCC signals with that of signals from nearby weld joint physical features, such as the weld crown, weld root and machined counterbores, which are ridges machined prior to welding to match unequal pipe wall thicknesses. Figure 2 illustrates the spatial relationship between an IGSCC and other geometrical reflectors in the vicinity. The photograph on top shows a weld metallograph of a field-removed specimen with IGSCC growing very close to the weld root and progressing into the weld. Indication location in the ultrasonic trace (or image) is one of the key considerations for discriminating IGSCC from geometrical reflectors. As shown in the figure, about 0.1- to 0.5 inch separates typical root, IGSCC and counterbore indications.

IGSCC DISCRIMINATION

Theoretical studies in the U.S. and U.K. have enabled IGSCC scattering models to predict responses for realistic inspection conditions (3,4). These have motivated the development of advanced signal processing methods that examine the signal temporal and spatial behavior to provide "features" to discriminate IGSCC from other reflectors (5). Field trials have been conducted to evaluate advanced, feature-based approaches for BWR weld examination under realistic plant outage conditions (6). Destructive tests are underway to compare with NDE data.

The EPRI NDE Center undertook the development of an expert system to integrate feature-based approaches with special knowledge used by experienced operators. An expert system shell program operating on a personal computer was chosen to codify the knowledge. To interpret the ultrasonic image, some key parameters that were identified are described below.

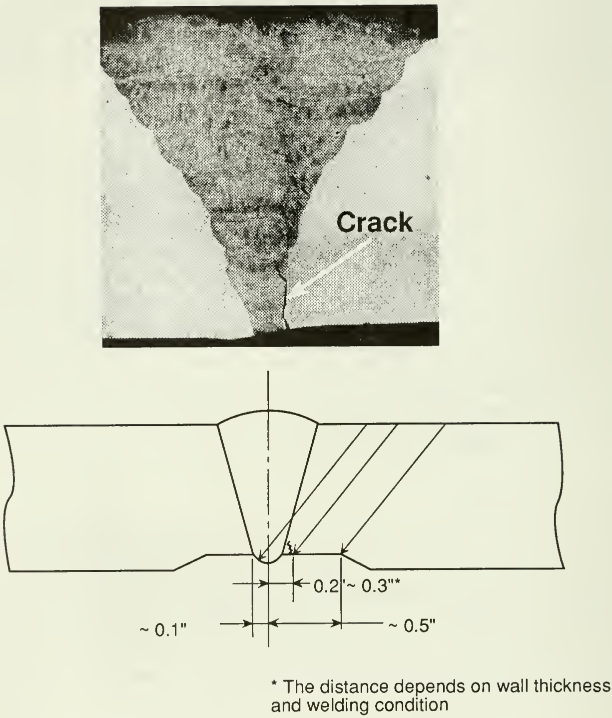


Figure 2. Sectional view of Pipe weld showing typical IGSCC and geometrical refelctor locations

Signal Amplitude

While signal amplitude is the primary means for detecting indications -- code guidelines require recording and reporting indications whose amplitudes are above established thresholds -- it is a poor discriminator of reflector type. There have been examples where signal amplitudes measured at different inspection angles were used to discriminate reflector types (7); however, they are not reliable discriminants.

Indication Location

Location is one of the key considerations for discriminating IGSCC, based on the reflector spatial relationship. Figure 3 is an example B-scan image presentation, the cross-section view, of a weld specimen similar to that in Figure 2. The B-scan clearly shows the counterbore, IGSCC and root image areas. The counterbore image is axially well separated from the crack and root images.

In many field welds, however, it is likely that the counterbore could be closer into the weld because of previous weld repair. Indication location may not be a reliable discriminator for such cases.

Metal Path

The distance along the beam axis is another essential parameter used to identify IGSCC and root signals. As can be seen in the B-scan image in Figure 3, the root signals occur later in time (hence metal path). However, counterbore indications sometimes occur at about the same metal path distance as IGSCC and cannot be separated, especially if the counterbore axial position is close to the weld root.

Amplitude and Arrival Time Consistency

Since counterbores and roots are machine-made reflectors, they are likely to be consistent in signal amplitude and constant in arrival time as they are scanned circumferentially. IGSCC indications, on the other hand, have different morphologies, follow grain boundaries and have facets. Their amplitudes are not expected to be consistent and their arrival times are expected to vary as they are scanned. It has been shown that spatial features related to amplitude and time-of-flight consistencies measured as a percentage of a standard were useful in making reliable separation (8). Figure 4 shows a scatter plot of these features measured for more than 50 reflectors, many of them field-removed samples of IGSCC and field-quality counterbores used to train and qualify personnel. The scatter plot shows the 95% confidence ellipse. It can be seen that these features are

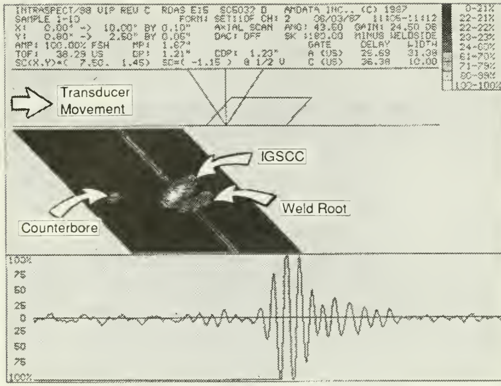


Figure 3. Example B- and C-scan presentations showing the axial separation between root and crack indications.

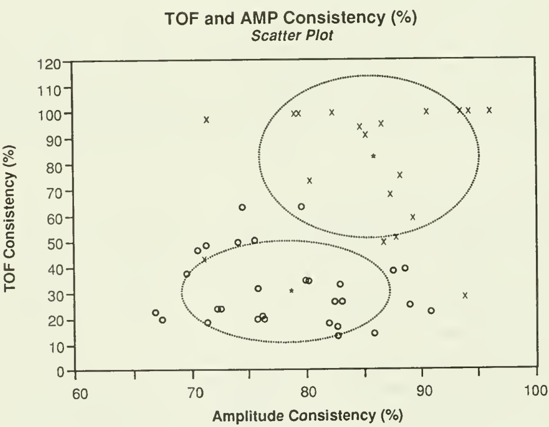


Figure 4. Scatter plot of spatial signal features for flaw discrimination.

reliable indicators; however, field-quality counterbores could be rough due to improper machining and could be confused with IGSCC.

Signal Echodynamics

The target-motion line, or the echodynamics, can reveal information about reflector type. Figure 5 shows echodynamics of different reflectors. The target-motion line for IGSCC tends to be straight and strong; and for weld roots it is expected to be "twisted" and wide. Small counterbores will have correspondingly short echodynamics; however, longer counterbores could appear similar to IGSCC.

Waveform

The characteristics of individual waveforms have been traditionally used by field operators. These include signal rise-time which tends to be short for IGSCC relative to weld roots.

Counterbore signals have several variations, depending on the machining quality. Figure 6 illustrates different examples.

Skewing the transducer in a plane parallel to pipe surface produces different responses. Counterbores and weld roots tend to persist for very small skew angles; IGSCC indicated tend to persist for large skew angles because of their facetted structure. However, for automatic systems skewing is difficult to apply because it requires a more complex mechanical scanner.

EXPERT SYSTEM FOR BWR WELD INSPECTION

Knowledge Base Development

The system consists of more than 300 rules in the knowledge base. Accumulation of the knowledge and encoding into the expert system shell to produce the first prototype was accomplished over a 6-month span (200 rules). This version was confined to consultation on the ultrasonic data only. The system was implemented on a commercial PC platform capable of controlling an automatic scanner around subject pipe-to-fitting component weld and digitally acquiring ultrasonic data. The rules were encoded in a question-answer format. The operator chooses the most appropriate answer that fits the data to questions posed by the system. The operator could invoke the feature-based imaging options during consultation to display and process B- and C-scans. Further, he/she could observe detailed signal behavior by invoking some of the signal processing options programmed into the

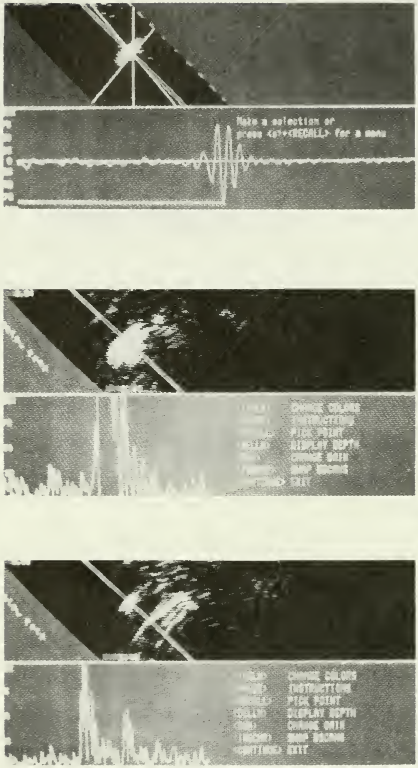


Figure 5. Example of echodynamic lines in a B-scan image. The top image shows the echodynamics for an IGSCC; the image in the middle is for an IGSCC close to the weld root, and the third image is of a counterbore and root. Weld roots have a wide and twisting lines and IGSCC lines are strong and straight.

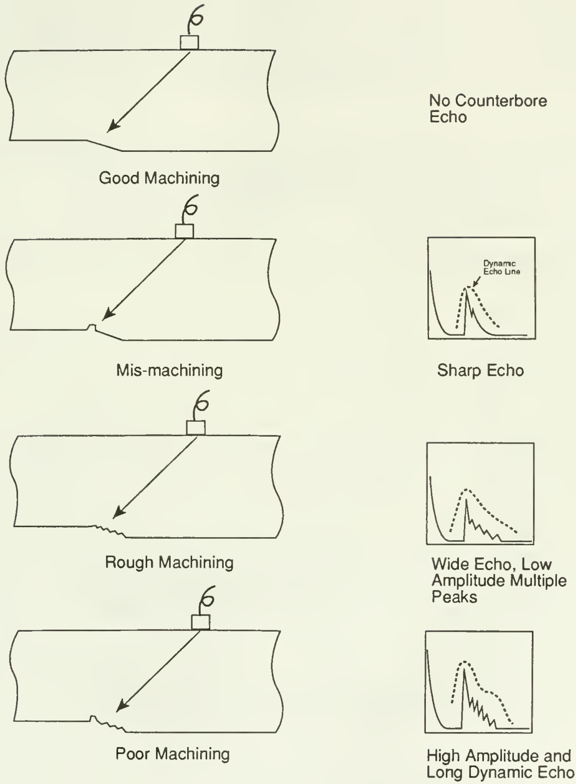


Figure 6. Examples of various counterbore conditions

software package. These include behavior of signal rise time, fall time, spectral content, amplitude and time-of-flight consistency measures, etc.

The historical rules were derived from interviews conducted among NDE Center staff members. The number of questions was limited to component age; inspection history of the component in question as well as other, similar component welds in the same plant and in sister units; and component material and configuration. Example rules are displayed in Figures 7(a) and 7(b).

Figure 7(a) shows an example of historical rules. Example 1 shows when favorable conditions exist for cracking. If the component is

- more than 10 years old,
- similar components in sister unit as well as in this unit showed evidence of cracking,
- past inspection revealed cracking,
- the material is stainless steel 304 material and the configuration was an elbow-to-pipe joint.

Then the most favorable condition for cracking occurs: this evidence is indicated as being close to 80%. The different "objects" relevant to UT IGSCC discrimination were: "location," "signal distribution," "multiple peaks," "echodynamic," "signal rise time," "echo front," "indication length," and "gate position." The relationship between these objects and reflector type were encoded, and rules to manipulate these were derived. The expert system was structured so that it confidently determined the possible reflector type solely from the indication location. It then methodically gathered auxiliary information to reinforce that decision; if such information were not present in the ultrasonic data it would "gracefully" fail to make a strong decision. Figure 7(b) illustrates two example rules. Example 1 is a simple rule that makes several interim conclusions on possible reflector types based on whether the time-of-flight locations map into the weld region. These conclusions include that the reflector is guessed to be a weld root with certainty 80%, a crack with 40%, etc. Certainty factors pertain to beliefs and vary from +100%, certain belief, to -100%, certain disbelief. Example 1 concludes that if the time-of-flight location is in the weld, the possibility of reflector being counterbore is -75%: counterbores are not machined in the weld. There is not complete disbelief (-100%), however, because the ultrasonic time-of-flight evidence may be faulty due to possible beam redirection at the weld fusion line. Example 2 considers a more complex rule based on signal distributions and behavior.

Weld History Rules

- **Example 1**

  If *Component Age = 10 (or more)* and *Cracking in Sister Unit and Welds in Similar Component = Cracked* and *Past Inspection = Cracked* and *Stainless Steel = SS304* and *Configuration = Pipe-to-Elbow*
  Then *History = Crack cf 80*

(a) Historical Data

Example Rules for UT Data

- **Example 1**

  If *Time-of-Flight = In-Weld*, Then *Guess-Root cf 80 and Guess-Other cf 60 and Guess-Crack cf 40 and Guess-Counterbore cf -75*

- **Example 2**

  If *Guess-Root and Distribution = Small and Indication = Long and Peak-multiple and Echo-dynamic = Wide*,
  Then *Signal = Root*

(b) NDT Data

Figure 7. Example rules used in BWR weld examination expert system

The UT decision was combined with available radiographic testing. Rules were developed to emulate operators in integrating the data. One of the factors considered was positive evidence in weld radiographs in influencing the overall decision; for example, the presence of geometrical reflectors in the radiograph could influence reflector decision based on UT. Similarly, if the UT decision was counterbore, the time-of-flight location was in the HAZ and the RT results indicated no reflector, then the combined decision weakened the UT counterbore decision.

System Evaluation

Figure 8 shows the circumferential area with the weld centerline (WCL) at the middle. Each 1-inch cell (or grading unit) which is exposed for examination (shown in white, the area not exposed for examination in dark) on both sides of the WCL with the reflector-type was marked with the system call.

For the purpose of evaluating the system, a technique was adopted to measure the number of correct and false calls. The crack detection rate, which is the number of grading units called cracked divided by the total number of cracked grading units, was defined. The false call rate was computed as the number of non-cracked grading units called crack divided by the number of non-cracked grading units. Both measures allowed for a one-grading unit tolerance, i.e., incorrect crack calls immediately adjacent to the correct crack cells are not accounted for in the false calls; nor are adjacent missed crack calls.

Figure 9 shows an example crack and the recorded crack calls ("C"). Four (4) of the six possible crack grading units were correctly detected by the candidate; therefore, the correct detection rate according to the defined guidelines is 67% (4/6). Of the other six uncracked grading units, two were incorrectly called cracks. However, one of the incorrect calls is adjacent to the crack and is within the one-grading unit tolerance. The false call rate is therefore 1/5, or 20%.

The procedure is similar to the means adopted in a Coordination Plan developed between the EPRI, NRC, and the BWR Owner's Group (9).

One of the authors evaluated the system on the inventory of field-removed samples at the Center. The data were previously acquired by a vendor; however, the results were not known. Based on above-described procedure for determining the performance, the correct detection rate was computed to be ~99% (69 out of 70 grading units), and the false-call rate was 7% (8 non-cracked units called crack out of 118). While the score was satisfactory, there was reason to suspect that intimate knowledge of the questioning "style" may have inherently biased the responses.

On an independent evaluation by another staff member the correct detection rate dropped dramatically: it was 12% correct detection rate with 33% false alarm. This difference in performance was attributed to the difference in familiarity with the questions and the questioning style. Several modifications were recommended to improve acceptability; some of them included rules that factored in weld history. These rules pertained to component operation time, weld type and location, past remedial repairs performed, whether stress relief procedures were applied in the past and changes, if any, in the water chemistry. It was also noted that some of the answers, especially in the UT questions, relied on qualitative answers for which the user required guidance. How wide is "wide" in the correct answer for echodynamics? How long is "long" for the indication length? It was decided to include screen help capabilities which provide examples and intent of the questions.

This revised system is being further evaluated at the NDE Center. It will also be evaluated by three utilities and a vendor. The purpose of these evaluations is not to demonstrate system performance; instead, the main purpose is to determine functionality of the system, accuracy of questions asked, need for additional questions and approaches for integrating additional knowledge and rules.

SUMMARY AND CONCLUSIONS

An expert system for assistance in interpretation of NDE data from boiling-water reactor welds has been developed on a PC system. A PC-based shell program was used to encode rules to discriminate intergranular stress corrosion cracking in BWR welds from benign, geometrical weld reflectors. The system has been integrated in a PC platform capable of automatic scanning and digitally acquiring ultrasonic data, and of imaging and feature-based processing. The expert system consists of approximately 300 rules. These rules include weld history and data from ultrasonic and radiographic testing. The rules for combining weld history information are less comprehensive than those for UT and RT data. The UT rules include specific temporal and spatial signal behavior that are automatically computed by feature-based imaging. The expert system combines results from ultrasonic and weld radiograph results to arrive at an overall decision on reflector type.

Data Sheet for Recording Results

Sample

Specimen ID: *B-4*

Inspected From *Pipe* Side

Inspected Area: *White Region*

Date:

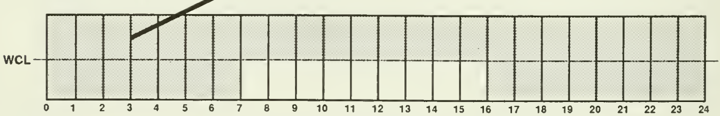
Team:

Test Mode: ( ) Manual / ( ) Auto

Instrument/System:

Inspector(s):

Safe-end Side



Pipe Side

Notes:
(1) Position on Outside Surface
(2) Each Block: 1.0-Inch Wide

Transducer

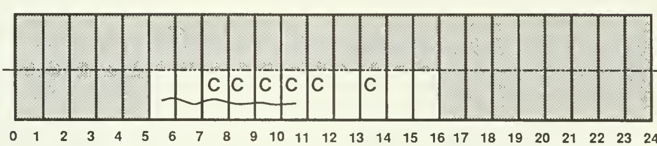
Type:
Frequency:
Element Size:
Focal Length:
Beam Angle:

Figure 8. Data Sheet for Recording Results

$$\text{Crack detectability (CD)} = \frac{\text{Units called cracked}}{\text{Total cracked units}} \times 100$$

$$\text{False call (FC)} = \frac{\text{Non-cracked units called cracked}}{\text{Total non-cracked units}} \times 100$$

Example:



$$CD = \frac{4}{6} \times 100 = 67\%$$

$$FC = \frac{1}{5} \times 100 = 20\%$$

Figure 9. Example data sheet and computation of correct detection and false call rates.

A preliminary evaluation on field-removed pipe weld samples with service-induced cracking revealed that the user had to be intimately familiar with the questioning style. The system was revised extensively to include on-line assistance to aid the user in answer selection.

The system is currently being evaluated at three utilities and at a vendor site, as well as at the NDE Center.

ACKNOWLEDGEMENTS

The work was supported by the Electric Power Research Institute, Palo Alto, California, under EPRI RP 1570-2. Dr. Gary Dau is the program manager.

The authors wish to express their sincere appreciation to Toshi Sasahara, Ishikawajima-Harima Heavy Industries (IHI), Yokohama, Japan, who shared in many discussions and documented his expertise while serving as a Visiting Engineer at the EPRI NDE Center. Melinda Harrell and Katie Estes, EPRI NDE Center, prepared the manuscript and the art work.

References

1. G. Dau, "Background to Recent U.S. Pipe Inspection Work," in Nondestructive Evaluation for Pressurized Components, London; Elsevier, 1984.

2. Automated Ultrasonic Pipe Examination and Data Interpretation. Palo Alto, CA: Electric Power Research Institute. EPRI NP-5760, April 1988.

3. Ultrasonic Scattering from Intergranular Stress Corrosion Cracks: Derivation and Application of Theory. Palo Alto, CA: Electric Power Research Institute, EPRI NP-3216, 1983.

4. J. M. Coffey, "The Reliability of Ultrasonic Inspection: Developments in Quantitative NDE in the United Kingdom," in Review of Progress in Quantitative Nondestructive Evaluation, Vol. 3, D. O. Thompson and D. E. Chimenti, Eds., New York: Plenum Press, 1984, pp. 1-12.

5. Nondestructive Evaluation Program: Progress in 1986. Palo Alto, CA: Electric Power Research Institute, EPRI NP-4902-SR, 1987.

6. Feature-Enhanced-Imaging Field Trials: Peach Bottom Unit 3. Palo Alto, CA: Electric Power Research Institute, EPRI NP-5865, 1988.

7. UT Operator Training for the Detection of Intergranular Stress Corrosion Cracking. Charlotte, NC: EPRI NDE Center, June 1983.

8. Spatial Features for Flaw Discrimination, Palo Alto, CA: Electric Power Research Institute, EPRI NP-5270-LD, 1987.

9. "Coordination Plan for NRC/EPRI/BWROG Training and Qualification Activities of NDE personnel." Rev. 0, January 1984.

ELECTRICAL SYSTEMS APPLICATIONS

Communications Alarm Processor

KEN HEMMELMAN, STAN BORYS, and JIM GRAFFY
Bonneville Power Administration
P.O. Box 491
Vancouver, Washington 98666, USA

RICK GOELTZ, STEVE PURUCKER, and BRUCE TONN
Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831, USA

ABSTRACT

In 1986, the Bonneville Power Administration (BPA) began a research and development project to build an expert system to analyze communications system and equipment problems. This project became known as the Communications Alarm Processor or CAP. The development of the CAP Project was contracted to DOE's Oak Ridge National Laboratory (ORNL) for development. The prototype was delivered in January 1989 for evaluation.

The CAP System has four primary goals:

1. Analyze operational communications system problems.

2. Reduce the bulk of raw data from the communications system alarm systems.

3. Provide statistical information about equipment performance with the goal of enhancing system performance and reducing the maintenance resources required to provide for acceptable system performance.

4. To give us some experience with expert systems in a control center environment.

BACKGROUND

BPA's telecommunications system is an integral part of the power system. We rely on the operational communications facilities to support stability control functions, high speed relaying (microwave transfer trip), SCADA control, various telemetering and data acquisition systems, and voice communications. We have 183 sites where high density microwave provides critical communications. There are 137 substations on SCADA control, 605 terminals of microwave transfer trip, hundreds of telemetering quantities, etc., that rely on our backbone telecommunications network. (See Figure (1), BPA Operational Telecommunications System.)

We have two systems specifically designed to monitor our communications systems and equipment to ensure reliable operation in support of the power system. These are the microwave alarm system (Badger), which reports on specific equipment failures, and the Microwave Monitor System (MWM), which is a real-time monitor of microwave system performance. The Badger tends to produce large quantities of data that must be interpreted by human experts to analyze equipment problems. Because of system requirements, some of the data is not standard. The MWM System does not produce large quantities of data, but the data is not very selective for isolating system problems. These systems do not readily provide for statistical analysis of the data. Special studies and/or data that is needed to evaluate various facets of system and equipment outages or performance must be done manually by human experts.

As we embarked on the development of this project, it was important to remember that our principal need was for "help" with the analysis of alarm data. Our first step in looking for the "help" was to look for technology that would provide a solution(s) to these problems. The fast growing field of expert systems seemed to provide these benefits, especially if we could marry an expert system to a good data base. This combination would provide for failure analysis as well as information concerning system and equipment performance. (See Figure (2), Basic Concepts of the Communications Alarm Processor.)

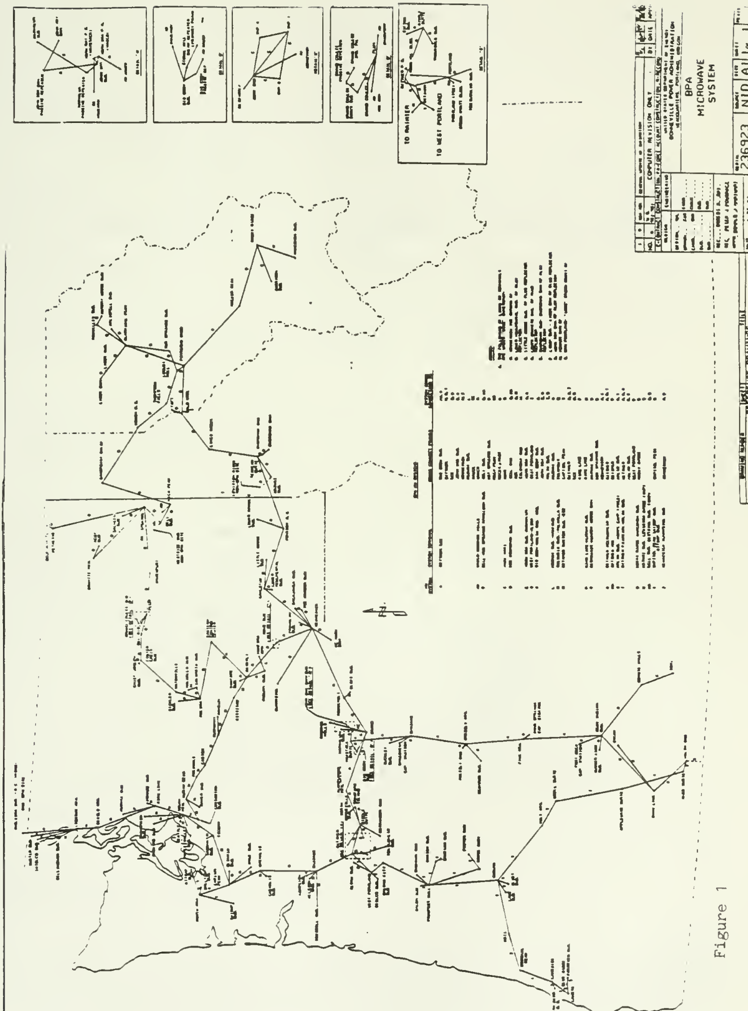

TO RAINIER
TO WEST PORTLAND
COMPUTER REVISION ONLY
BPA
MICROWAVE
SYSTEM
236923
NID
A1
REFERENCE DRAWINGS

Figure 1

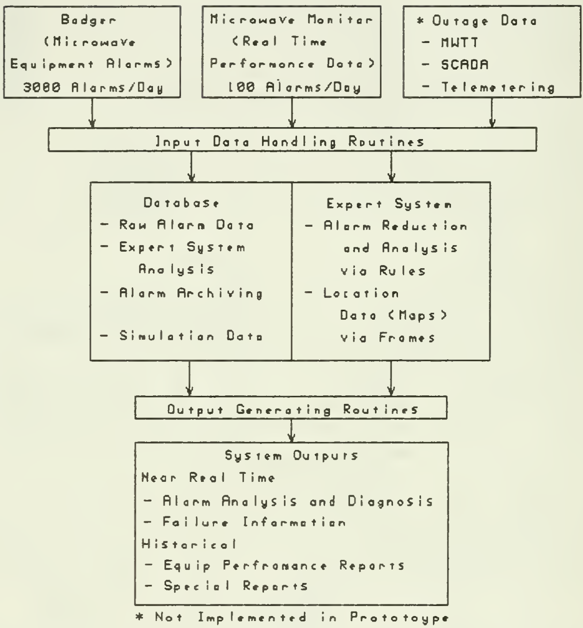


Figure 2. Basic CAP Concepts

We developed some of the basic concepts for the project in house. To verify the conclusions we had reached, we contracted with ORNL to do a study of our situation. They concurred that this approach would be very suitable. ORNL made a study of the expert system shells that were available and the data bases that would meet our needs. They also looked at the hardware requirements that we would need to implement the system.

As part of the preliminary study, we asked ORNL for recommendations on the feasibility of implementing the entire system as we had envisioned, or implementing a smaller prototype. Their recommendation was to implement a prototype using only one of the seven major microwave networks (the "N" System), and looking at only Badger and MWM data. This had the benefit of allowing us to evaluate a system, confirm the benefits, and ease some of the performance parameters of the system (primarily response time).

From their recommendations, we moved forward with the design of the project using the hardware and software that was proposed. We entered into a contract with ORNL to design and deliver the CAP System.

It is interesting to note that ORNL identified several research challenges that the CAP Project presented.

- Asynchronous input data
- Continuous operation
- Uncertain or missing data
- Expert System/Operator Interface
- High Performance
- Nonmonotonically
- Temporal reasoning
- Focus of attention
- Integration with procedural components
- Guaranteed response time

PROJECTED BENEFITS OF THE CAP PROJECT

In the beginning as we analyzed where we were, what our needs were, and where we wanted to be with the alarm summaries and analysis, we identified potential technical benefits for the project. As with most utilities, we were and are under pressure from management to become more effective in the operation and maintenance of the power system and the supporting telecommunications equipment. Working towards that goal, we projected a set of benefits that the CAP Project would provide:

- The system would provide for near realtime (NRT) alarm analysis and data reduction. In times of major outages, operators are overwhelmed with alarms, most of which are "effect" alarms that hide the "cause" alarms. The system would help to alleviate this problem. There would be less need to have human experts available to analyze every system trouble, as well.

- We could readily analyze data to establish information about equipment performance. With this information, we could tailor our maintenance program to attack those areas where the need is greatest. Similarly we would not waste resources on equipment that is performing adequately.

- With the query capability of the statistical data base, our engineers could request varied information to help them operate and maintain the systems and equipment.

- It would allow BPA to gain experience in expert systems in the NRT environment of our operational control center. We recognized that there are many situations beyond CAP where there are potential benefits for the use of an expert system.

- It would give our design engineers an opportunity to work with the knowledge engineer from ORNL in order to gain experience for future development of expert systems at BPA.

- We would have the hardware and software to allow future development of expert systems for other applications.

SYSTEM DESCRIPTION

The CAP integrates an expert system, Nexpert Object, and a statistical data base, SAS, to form the basic system. It runs on a VAX Station 3200 with full graphics support. Input/output handlers are written in C to integrate the various software components. (See Figure (3), CAP Prototype System.)

The realtime alarm data is captured by the system and stored in input data buffers (IDB), one for Badger and one for MWM. In each case, the alarm message basically contains date/time, location, alarm message, and occur or clear. One of the major concerns with the system is the time factor. Alarms do not arrive at the CAP together, nor are they likely to arrive in the proper sequence. Because of the dynamics of the communication system, data may be relatively old and yet critical to an analysis.

The expert system provides for the analysis of the alarms. Within the expert system, the relationship of alarms and failures are handled with rules. The rules were developed from fault trees that were derived by the ORNL knowledge engineer as he interviewed BPA's human experts. The fault tree for a relatively simple condition, excessive phase jitter, is shown in Figure (4).

Figure (5) shows fault trees for more sophisticated problems, Noise Outage and Noise Performance. There are many rules associated with the analysis of noise. With expert systems, it seems that someone always asks: "How many rules?" There are about 250 rules in the CAP System. Many more rules would have been required unless confidence factors, reflecting experts' judgment, were used.

Because several different alarm conditions could be in progress at different locations on the communication system simultaneously, BPA experts developed a list that prioritizes alarms for the expert system. At the top of the list is noise outage, which is most critical and the condition the expert system tries to diagnose first. There are 13 other alarm categories below this in a descending order.

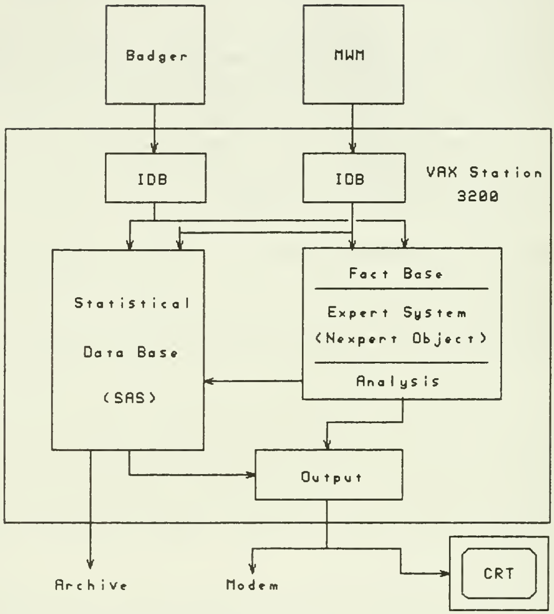
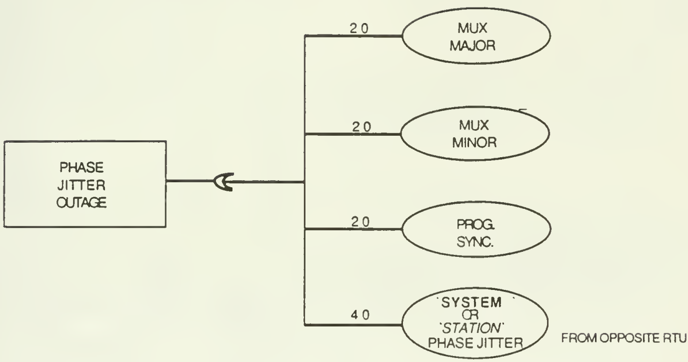


Figure 3. CAP Prototype System

PHASE JITTER OUTAGE 'SYSTEM' OR 'STATION'

K-PJO-8-9-88



NOTES:

1. THESE ALARMS ARE CAUSED BY STATION PROBLEMS IT IS NOT A HOP PROBLEM.
2. THIS ALARMS CAN EITHER BE A 'STATION' OR '**SYSTEM**' ALARM. THE '**SYSTEM**' ALARM IS PRESENT WHEN THE MWM CANNOT IDENTIFY SPECIFIC STATION CAUSING THE PHASE JITTER.
3. FOR '**SYSTEM**' OUTAGE ONLY-LOOK- AT DITT ONLY WHICH IS THE (SOURCE OF SYCN. PILOT) OTHER STATIONS WILL HAVE ALARMS, BUT FOR '**SYSTEM**' OUTAGE THESE ARE FALSE ALARMS.

7-22-88

Figure 4

NOISE OUTAGE AND NOISE PERFORMANCE FAULT TREE

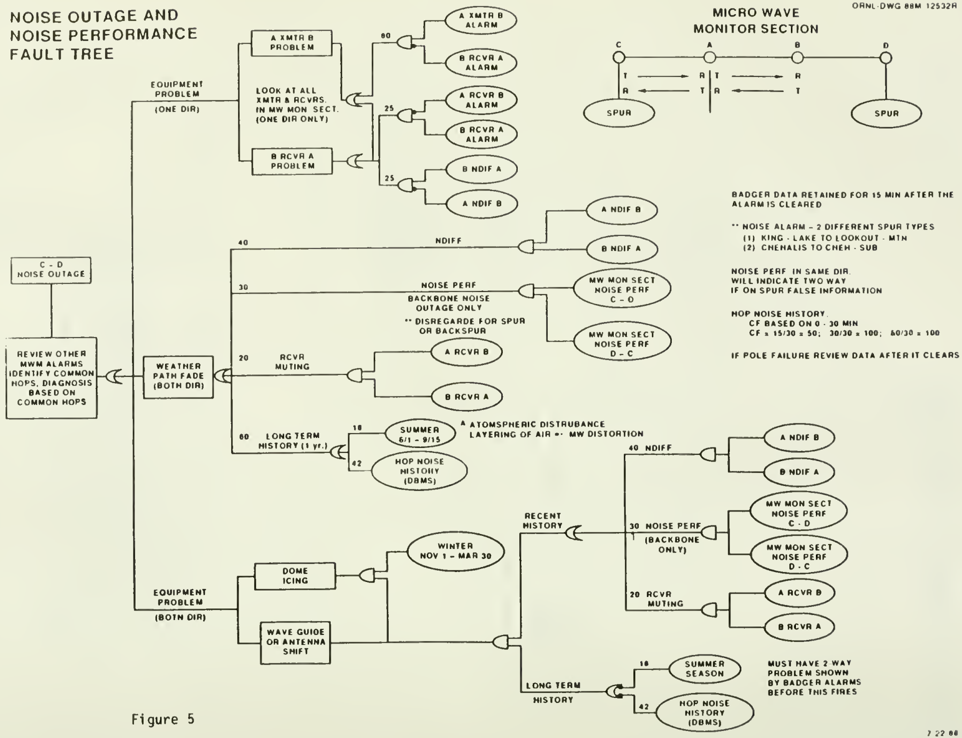


Figure 5

The interrelationship of locations (microwave sites, substations, etc.,) is handled with frames. Frames are ideal for this application as they possess a strong inheritance capabilities. (Figure (6) is an example of a frame.)

With this technique using rules and frames, the rules can be generic. The interrelationships of the alarms at various connected or unconnected sites can be readily resolved.

As each problem is analyzed, a "confidence factor" is calculated for the particular problem. It uses the formula:

$$CF(O)=[CF(a)/100+(CF(b)/100)((100-CF(a))/100)]*100$$

This is a form of the certainty factor rule where the certainty factor range is between zero and 100. Several alternative calculations were tested that did not fit our process. If you look at the fault trees of Figure (4) and Figure (5), you will see the confidence factors as numbers near the elipses.

Two classes of information are provided to the user by the system. The first is "near realtime" data. We specified in the requirements that we would like to have analysis of system problems within about 30 seconds of the event. Our experience in the control center environment indicated that waiting much longer makes the operators very nervous, and limits their "comfort" with the system. This placed a strong requirement on processing speed for the CAP.

The second class is historical data. The time requirement for this data is "within 24 hours." In general terms, historical information on equipment performance is not time critical. If a piece of equipment is showing abnormally high outage time indicating that maintenance is required, the scheduling of crews, etc., indicates that 24-hour response is acceptable. In practice, we may run this type of summary reports at midnight when system activity is typically low.

Failure information is presented to the user as a text display. It is prioritized with the most likely cause of the problem, as determined by the confidence factor, being presented first. The expert system may find several potential causes, that are presented to the user in descending order.

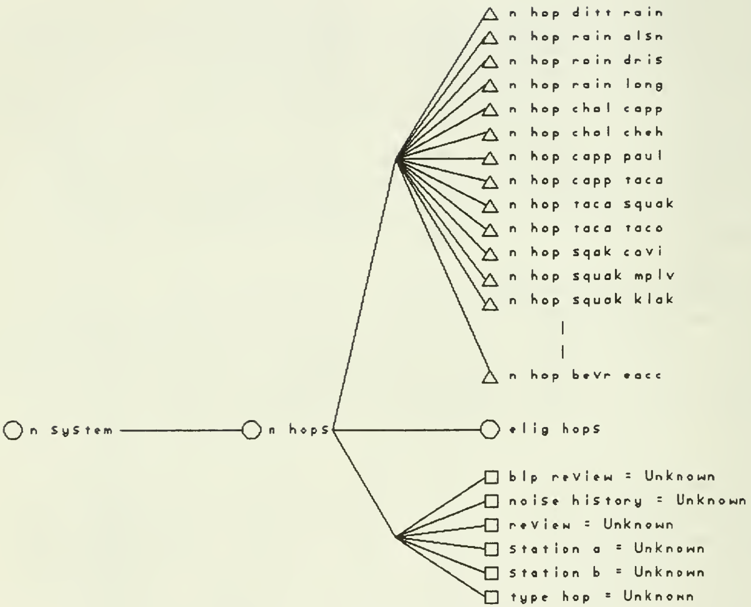


Figure 6. Sample Frame

The system also has a simulation mode. This provides the capability of using running an offline analysis with a specified set of alarms to verify that the analysis made by the expert system is correct. It also allows events on the system to be rerun through the expert system to confirm the diagnosis, or the lack of diagnosis.

Finally, the system provides for alarm archiving. If we continued to accumulate alarm data, soon our main memory would overflow. The system archives alarm data after it is no longer useful and has been verified (or corrected) by the operator. Data on alarm conditions that have not reoccurred within 15 minutes is no longer needed for diagnosis.

PROJECT STATUS AND INITIAL OPERATING EXPERIENCE

The CAP prototype was delivered by ORNL in late January 1989. It is installed in our Dittmer Control Center. We have begun to evaluate the performance of the CAP System. We are finding that there is a substantial learning curve in dealing with an expert system. It is different from the typical computer system that most of us, and most programmers, are familiar with. As we gain experience, our intent is to make a critical analysis of the application of expert systems as they apply to the near realtime situations on the power system.

Four days after the CAP was operational and the ORNL folks had left, the first significant problem occurred to the communications system. It was an unusual problem that had not been covered in the fault trees. (An impedance matching transformer that was associated with the baseband bridge failed.) While the CAP understandably misdiagnosed the problem, but it did correctly determine the location of the failure. Since that time, we have had several minor problems with the CAP System. A typical example is that the IDB for the MWM hangs up, but the IDB for the Badger works properly. We do not perceive these problems to be major, but they have limited the amount of experience we have had to date.

The ORNL staff is in the process of developing statistical analysis routines (using SAS) to analyze CAP alarms. Total amounts of alarm activity, both frequencies and durations of alarm occurrences, are used to identify potential microwave equipment problems. For example, the microwave stations engine generator (EG) runtime is an important maintenance item. CAP analysis accumulates runtime for each EG with a future consideration of doing maintenance on an "as required" basis. (See Figure (7), EG Runtime Summary, Simulated.) A second type of analysis technique uses standard deviations to identify equipment that are marginal performers. A third analysis technique compares performance measures that should have a predictable relationship. For example, noise differential is summarized for both directions of a path. If the ratio of the summaries indicates imbalance (i.e., the ratio is not close to 1.0), then a potential problem area has been identified.

We anticipate that the information we will get from the system will be very useful. One important aspect that the expert system plays in developing the data for the alarm summaries is that it identifies the cause of each problem. This is important in that it filters out the effect alarms. For example, if we are tracking receiver performance, we want to track only alarms that are caused the by a receiver failure. We do not not want to include receiver alarms that are the "effect" of a transmitter failure.

Again, with the analysis we plan to be able to direct our maintenance and to the most needed equipment. This has substantial potential in a time where resources (staff) are limited.

FUTURE ENHANCEMENTS

The CAP is a prototype system. We anticipate that over the next year, the CAP knowledge base will be validated and improved. Fault diagnostic logic will be refined and added according to real world operating experience.

In the immediate future, we plan to add a feature to improve the determination of confidence factors. As the communications system changes, the confidence factors that are used by the expert system need to change. For example, if during the winter a microwave antenna is damaged by ice, that path will likely see a decrease in signal and an increase in noise. We want to automatically adjust the confidence factor to take into account the degraded path, and analyze the path for other problems setting aside the path problem.

3-13-89

EG Runtime Summary

Location	Run Time		Minutes	Date	Notes
BEVR	1230	1945	435	11-28-88	
	0200	0939	450	12-18-88	
	0635	0845	130	1-22-89	
	0136	1455	799	1-31-89	
	1200	1725	325	2-14-89	
	0345	1955	970	2-17-89	
	1245	1325	40	2-24-89	
	0136		1356	3-01-89	
		1455	895	3-02-89	
Total			6253		100 Hour Sevrice Due
BEVR	0345	1945	360	3-09-89	
	0654	0802	68	3-11-89	
. . . .					
CHAL	0720	1900	740	12-07-88	
	0936	1045	69	1-28-89	
	0036	2057	1221	2-28-89	
. . . . .					

Figure 7. Simulated EG Runtime Summary

This enhancement will allow the expert system to look at historical data and automatically modify or update these confidence factors. This has some analogy to a "learning" system. As with the development of the original system, development of the modification is being done by ORNL.

We designed the prototype CAP to analyze the data from one of our seven major communications systems, the "N" System. The "N" System is our largest system, containing almost 1/4 of our microwave network. We plan to expand the prototype to encompass all of our major microwave systems. We are beginning to look at the capacity of the VAX Station 3200. It may be that we will need to add some parallel processing to keep system performance acceptable as the other microwave systems are added. It is too early at this time to make a judgment on this.

Another future enhancement will add a graphics display to the system for the display of the various diagnoses. We have historically used "maps," "block diagrams," etc., to display failure and outage information (such as power system status and information that the dispatcher sees). With an expert system, there is knowledge to be displayed that may be better conveyed with graphical displays. With the expert system, we determine alternate solutions to a problem. While some of these solutions may be less probable than the solutions originally presented to the user, they will in some circumstances be the correct solution. A good method of presenting this information needs to be developed and tried. We believe that a graphical display will be useful in the presentation.

CONCLUSION

The CAP Project is our first significant expert system development at BPA. While it is still in its infancy, it appears to have benefits for us. The marrying of the expert system with the statistical data base appears to be a step in the right direction in providing failure information and outage data to support our operation and maintenance activities.

The outputs that we feel are most important are:

- The diagnosis of problems on the communications system to the specific station.

- The identification of equipment that shows substandard performance.

We believe that the enhancements to our maintenance activities will in essence "pay" for the system. As time goes by, we will be able to evaluate the benefits of the expert system with more certainty. We believe that expert systems have applications in a control center environment.

REFERENCES

1. B. E. Tonn, R. T. Goeltz, and S. L. Purucker, "Expert Systems Microwave Communications Systems Alarms Processing: A Feasibility Study," Oak Ridge National Laboratory, ORNL/TM-10429, July 1987.

2. B. E. Tonn, R. T. Goeltz, and S. L. Purucker, Oak Ridge National Laboratory; T. Wiggen, Oak Ridge Associated Universities; and D. MacGregor, MacGregor-Bates, Inc., "Bonneville Power Administration Communications Alarm Processor Expert System: Phase I Description," Oak Ridge National Laboratory, ORNL Draft, February 1989.

A Generator Expert Monitoring System

B. LLOYD
Ontario Hydro
Toronto, Ontario, Canada

W. PARK
SRI International
Menlo Park, California, USA

J. WHITE and M. DIVAKARUNI
Electric Power Research Institute
3412 Hillview Avenue
Palo Alto, California 94304, USA

ABSTRACT : *The reliability of turbogenerators is critical to the overall reliability and operation of any power plant. With the current trend towards refurbishment and life extension of existing plants, the average age of generators is increasing. Thus it is becoming even more important to improve generator monitoring systems and to provide early warning of machine problems before failure and a prolonged plant outage can occur. Although considerable generator diagnostic information is often available, it is not always correlated or otherwise analyzed and presented in a form which can best be used by generator operators.*

This paper describes work currently underway on EPRI project RP2591-3 entitled "Generator Expert Monitoring System (GEMS)", to develop an on-line generator monitoring system using expert systems technology. This system will correlate generator diagnostic information from existing sensors to provide operations personnel with warning of developing generator problems and recommendations for corrective action. Developing the software for GEMS presents many technical challenges associated with the requirement for a real-time expert system which can be readily customized and applied to generators of varying design, manufacture, and operating environments. A description of the software architecture currently being implemented to meet these requirements is given.

INTRODUCTION

Monitoring systems for generators are used to warn of abnormal conditions developing in the machine before significant damage or failure can occur. A major insulation or core failure can result in a six month to one year outage costing several millions of dollars. Although such major failures are infrequent, other less catastrophic failures occur more frequently and the overall result is a less than satisfactory generator forced outage record. The fact that a significant proportion of any utilities' generating capacity is needed to provide for the unreliability of generators combined with the high cost of outages and repairs provides a very strong incentive to develop methods to obtain better performance and reliability from our existing plants.

Considerable generator diagnostic information is normally available. Examples include core monitor output; stator winding, cooling system and core temperatures; vibration of core, frame, bearings and endwindings; etc. Also, considerable information is available from the auxiliary process systems of generators (for example water, oil and excitation). Although this information is more or less readily available, it generally is not correlated or otherwise analyzed and presented in a form which can be used by operations personnel. The objective of the GEMS project is to develop an on-line generator monitoring system using expert systems technology. Expert system techniques have been used in many applications [1,2,3,4] and offer the opportunity for significant improvement in generator monitoring systems.

Two key requirements in the design of GEMS are described in this paper. Software techniques to obtain the real time processing capability necessary for monitoring turbogenerators and techniques for easily customizing and tailoring the expert system for a particular generator configuration are outlined. Software for the prototype monitoring system is currently under development. A prototype framework has been completed and specific reasoning covering some generator subsystems has been encoded. The first installation of this system on an operating generator will be made in May of 1989.

SYSTEM DESCRIPTION

Capabilities

The expert monitoring system will use data input from available sensors (or sensors that could be easily and economically retrofitted to the generator) to provide an on-line monitoring tool to assess generator condition. Turbogenerator operators and their supervisors are responsible for evaluating the generator status and if problems arise, taking the necessary corrective action to bring the generator back within safe operating limits. In general, operators only become aware of developing generator problems when a sensor alarm threshold has been reached. At this time, the operator must assess the status of the machine from the available sensor indications and make a decision as to the course of action required to further diagnose or remedy the problem. Often this decision is made under tight time constraints and is based on a limited amount of uncorrelated information of sometimes dubious accuracy. Additional checks or generator maneuvering may also be required before the alarm can be verified and corrective action taken. In practice the generator is often allowed to run until it automatically trips as a result of winding failure, fire, etc. The goal of GEMS is to improve this situation by continually monitoring and correlating sensor data and providing operations personnel with reliable advice on corrective action when a problem is detected.

As an example of the capabilities provided by GEMS, consider the example of a single stator bar blockage in a direct water-cooled generator with and without GEMS. Using traditional monitoring techniques, the operator would probably not become aware of the problem until the coolant hose outlet temperature alarm limit was exceeded for the particular blocked stator bar (assuming that all stator hose outlet temperatures are continuously monitored). Normally this alarm level would be set significantly beyond the nominal temperature for the coolant hose outlet under full load conditions. If the generator was operating at reduced load, this alarm (and any warning to the operator) would only appear after a very serious condition had existed in the machine for a significant period of time. A temperature alarm could result from problems within the machine that fall into three general categories; instrument error, overloading, or inadequate cooling of the stator winding. The operator would have to manually check the status of all slot temperatures, all outlet hose coolant temperatures, coolant flows and pressures, coolant inlet and outlet bulk temperatures, phase currents, core monitor output, excitation level etc. Before diagnosing the problem as a blocked cooling passage in a particular bar, the operator must consider and eliminate many other potential problems that would result in the same alarm. He must be fully aware of all these other problems and their impact on the generator, have enough time to complete checks on various other sensors, and be able to interpret a large amount of data which in some cases may be incomplete or inconsistent due to sensor failure etc. This requires a great deal of judgement under considerable pressure. Assuming the operator has analyzed the situation correctly, he is now faced with a decision as to the correct course of action to alleviate the problem and restore the generator to a safe operating condition as quickly as possible. As described in this scenario, monitoring is currently based on general alarms, relies entirely on the operators experience, does not provide early warning of developing generator problems, and leaves considerable room for error in the detection, diagnosis and correction of generator problems.

Considering the same scenario described above with a GEMS installed, the operator would receive much earlier and more specific warning of the overheating condition allowing time for appropriate corrective action to be taken. GEMS would be continually monitoring and correlating all the available generator sensors. On a continuous basis GEMS would scan and check for abnormalities in slot temperatures, coolant outlet hose temperatures, coolant pump status, coolant flow and pressures, hydrogen temperatures etc. In many cases, the alarm levels for GEMS are calculated dynamically as a function of generator load or other operating conditions. Thus GEMS is very sensitive to small deviations in sensor behaviour, long before a serious condition has developed. Once a small abnormality in a particular hose coolant outlet temperature was detected, GEMS would use other relevant sensor data to analyze possible causes for this condition. Problems such as sustained overload (failed AVR), loss of coolant, high winding current, broken strands, etc would be considered by GEMS and compared to the current state of the generator sensors. The operator would then be provided with a list of one or more suspected problems that are consistent with all other sensor indications. In this case, GEMS would report a high probability of a blocked stator bar with the explanation that this conclusion was based on a rapid rise in a particular outlet hose temperature, a slot temperature for this bar rising, slot temperatures for adjacent slots rising, and other sensors in the cooling and stator winding systems remaining normal. GEMS would also provide suggestions for operator corrective action. In this example, the operator would be advised to do a fast unload on the machine, maneuver at low load to confirm the blocked stator bar, and then shut down for repair.

An incident similar to this occured at an Ontario Hydro Nuclear generating station. On this 500 MW unit, all generator stator temperatures are continuously monitored by a sophisticated on-line monitoring system called a generator temperature monitor (GTM). The GTM uses algorithms to calculate dynamic temperature alarm limits as a function of generator loading. During a recent run-up after a maintenance outage, the GTM alarmed on high stator bar temperatures. Although the temperature was not above the high limit alarm (90C), a number of stator bars had temperatures exceeding the dynamic alarm limit for the low load conditions. Had there been no real-time, on-line, dynamic monitoring of the stator temperatures, the machine could have severely overheated resulting in an outage of several months to replace the overheated bars. Even with the GTM system in place, it required about a day and a half to verify the alarm and determine where the blockage was in the stator cooling system. Had GEMS been used on the unit, a clearer indication of the problem and its location could have been provided immediately resulting in an additional saving in the day and a half outage time on the nuclear unit. Thus even in the case where a fairly sophisticated alarm system is in place, it may be possible to justify GEMS on the basis of the incremental saving in identifying and locating generator failures.

Real-Time Operation

A key benefit of GEMS is the ability to provide warning of developing generator problems before maximum sensor limits are reached so as to limit the extent of damage to the machine and give operators sufficient time to take corrective action. In order to provide this capability, GEMS must be continuously sampling and analyzing all sensor data in as short a time frame as possible. Depending on the generator design, readings from as many as 300 individual sensors may have to be evaluated. The time taken by GEMS to cycle through and analyze all this sensor data must be faster than the time required for most serious generator problems to develop. A maximum cycle time for GEMS has been established at 3 minutes. The types of problems GEMS will detect are those which occur with sufficient warning time to allow corrective operator action and can be detected without resorting to specialized sensor technologies. A partial list of typical problems detected by GEMS is given in Table 1.

Both swiftly developing problems and problems which develop over a long time frame are difficult to detect. In the case of a swiftly developing problem, for example a wiped bearing, no early warning to the operator may be possible. Conversely, because it is necessary to ensure a response time for GEMS on the order of several minutes, it is impractical to store and reevaluate a mass of long term sensor data searching for slowly changing sensor deviations. Thus, very long term generator problems may not be recognized until significant sensor deviations have occurred. Therefore a compromise is necessary for the processing speed and problems GEMS is designed to detect. The approaches selected for use in GEMS to attain practical data processing rates are discussed in the section on software architecture.

TABLE 1
Typical Problems Detected by GEMS

- Reduced cooling flow in the stator winding
- Unbalanced current in winding parallels
- Phase unbalance
- Sustained overload
- AVR malfunction
- Exciter power stage fault
- Hydrogen cooler blockage
- Rotor thermal unbalance
- Poor rotor shaft grounding
- Transient induced core burning

Adaptability

To be useful to as many utilities as possible, GEMS must cover a range of turbogenerator manufacturers, sizes, and configurations. Most utilities have generators from two or more manufacturers. These machines may have two or four poles, have a variety of ages, and employ greatly different numbers and types of sensors. There can also be differences in operating practices from utility to utility or even from plant to plant. The cost and difficulty of customizing GEMS for a given installation must be kept to a minimum. Major software revisions for each installation would result in an impractical and expensive GEMS. Thus the GEMS software must be designed to be easily adapted for use on different generator types and configurations. As part of the GEMS software development, a separate Installation Advisor program will be developed to lead utilities through the steps to configuring the expert knowledge base. The Installation Advisor program will allow individuals knowledgeable about turbogenerators to configure the GEMS software for a particular site.

Generator instrumentation is normally provided by the generator manufacturer and can vary significantly with the size, age, and type of generator. During the GEMS installation, factors such as the number of sensors, sensor types, sensor locations, etc will have to be customized. Other factors such as normal operating points and alert thresholds will also have to be determined. This information is required so that GEMS can reason with the sensor data and provide clear advice on the location, urgency, and severity of a problem. Physical information about the various generator components and their layout is also necessary. For example, when considering the stator winding, GEMS will have to have information on the number of parallels in the winding, the number of slots in the core, and the location of each bar (top or bottom of the slot) in the winding. For other systems, such as the auxiliary cooling systems, GEMS will have to know the interconnection details and the location of various pumps, valves and filters.

As well as providing flexibility in specifying the configuration parameters for a particular site, the Installation Advisor must also allow flexibility in the type of advice that GEMS will provide for specific generator problems. The advice from GEMS must not conflict with the operating policies and procedures in place for that particular unit (for example, the criteria for reducing load on a baseloaded unit may be different than that for a peaking unit). During the GEMS installation all of these parameters will have to be examined and specified for the particular unit of interest.

The Installation Advisor program is critical to the commercial application of GEMS. GEMS must be built with a high degree of flexibility, thereby limiting the cost of installing and tailoring the software for a particular site. A large portion of the knowledge engineering task for GEMS has involved identifying areas where the knowledge base will have to be made flexible and means for obtaining this flexibility. GEMS is structured to contain a generic model of a generator which can then be customized by pulling in specific information for a particular configuration. The configuration process is menu driven and does not require knowledge of the GEMS software architecture or software programming techniques. Modifications made using the Installation Advisor program do not affect the basic reasoning core of GEMS, but involve creating configuration files containing specific site information. The software architecture to facilitate this flexibility is discussed in the next section.

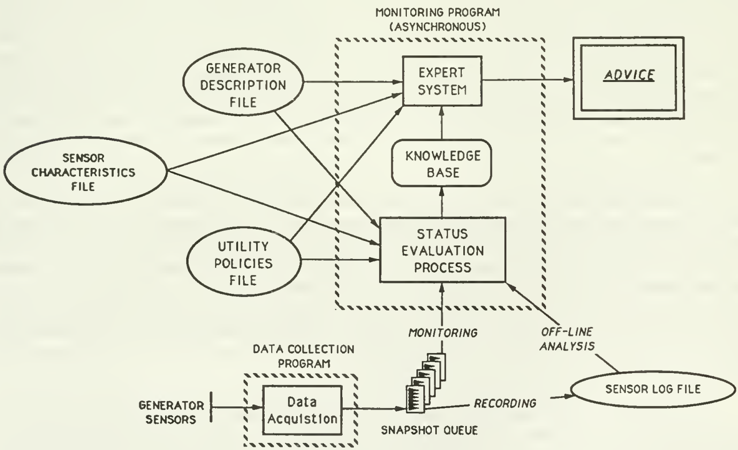
SOFTWARE DESCRIPTION

GEMS software consists of two independent program modules; the intelligent Monitoring Program (expert system) and the Installation Advisor program used to customize the Monitoring Program for a particular generator site. Both programs are being written in a commercial expert system shell (Automated Reasoning Tool - ART- from Inference Corporation). A number of large expert system shell programs were evaluated for use in this application. The ART shell was selected because it provides many useful knowledge representation schemes while still maintaining relatively fast rule processing speeds.

Monitoring Program

The expert system software for GEMS resides in the main monitoring program. This program evaluates sensor data and provides operators with actionable advice based on sensor deviations. The monitoring program is divided into two subprograms; one component which can be best described as the Expert System part of GEMS and another program called the Status Evaluation Process (Figure 1).

Figure 1
GEMS
Architecture

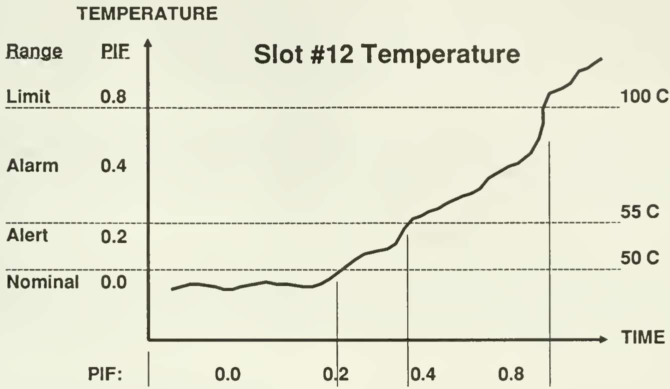


The Status Evaluation Process is a fast procedural program which identifies and classifies abnormal sensor indications for evaluation by the expert system. By off-loading the mathematically intense procedural software from the Expert System, GEMS can be run on a much smaller computer and still maintain an acceptable real time response. The Status Evaluation Process is written in Common Lisp. Using information from the Generator Description File about the particular sensors in this generator, the Status Evaluation Process produces a set of facts about the generators current status for use by the Expert System program. Each sensor reading from the generator is quantized into one of four possible ranges; nominal, alert, alarm, or limit. Thresholds for these ranges are established for each particular generator by the Installation Advisor Program. In many cases, the ranges for a particular sensor may be calculated as a function of some other sensor values. For example, ranges for the generator stator winding temperatures are a function of the bulk coolant inlet temperature and the stator current. The Status Evaluation Process also computes trends for each sensor reading and quantizes these into rising, falling, or steady. In some cases, ranges and trends may also be calculated for a predefined group of sensors to form a more complex indication. For example, temperatures from each phase of the stator winding are averaged and compared with each other as well as with valid ranges calculated from the temperatures for each stator bar in the phase. Once a complete snapshot of sensor data for the generator is evaluated by the Status Evaluation Process, facts about the quantized sensor ranges are asserted in the current fact database for interpretation by the Expert System program. Each data snapshot is treated independently except for sensor trends calculated by the Status Evaluation Process.

The Expert System portion of the monitoring program evaluates the information produced by the Status Evaluation Process to produce a list of possible generator problems. In many cases, this may require physical information about the generator design (which would be obtained through the Installation Advisor program) or the correlation of sensor indications from various dependent generator subsystems. For each problem diagnosis a certainty factor is calculated. This certainty factor is based on the range and trend of the currently evaluated sensor data snapshot. The ability to provide an estimate of the confidence in a diagnosis based on the current sensor indications is an important aspect of GEMS. In the early stages of a developing generator problem, the sensor indications may be ambiguous and a large number of possible problems may be suspected. GEMS must therefore provide to the operator some indication of the most likely diagnosis. As the problem worsens, sensor indications will deviate more from normal, and the confidence for a small group of problems (or only one) will increase while confidence in other diagnosis will decrease.

A number of different approaches for implementing confidence calculations were considered for GEMS. The approach selected is a hybrid of several more complex techniques. The particular approach selected for GEMS has the advantage of not requiring a huge amount of computing resources for calculating confidence while still having enough depth so as to match the level of complexity in the knowledge base. Because of its simplicity, the approach selected for GEMS is also understandable for the generator experts who are designing the knowledge base. Experts in machine diagnosis weigh each problem indication according to both its magnitude or strength and to the specificity of the indication to the problem being considered. To mimic this mode of reasoning, GEMS computes the net confidence in a particular problem diagnosis by multiplying together two weighting factors.

The first factor, called the Problem-Indpendent Factor (PIF), allows GEMS to take into account the strength of a problem indication. The PIF increases from zero to one in discrete steps as the sensor indication deviates farther from its nominal calculated range. For example, Figure 2 shows the temperature of slot #12 in the stator winding of a generator. In this Figure, the temperature starts out in its nominal range, which does not indicate any problem, so the initial PIF for this indication is zero. As the temperature begins to rise, perhaps due to a blockage to the coolant flow in one or both of the bars in that particular slot, the PIF is increased in increments. All sensor readings are divided into 4 ranges, normal, alert, alarm and limit, with higher ranges resulting in a larger PIF. Therefore as a sensor moves into a higher range, the belief in a particular problem (or group of problems) indicated by that sensor increases.



The Installation Advisor program can be used to specify the PIF value for each indication range, or the following default values can be used:

Indication Range	PIF Value
Limit	0.8
Alarm	0.4
Alert	0.2
Nominal	0.0

The second factor for confidence calculations, called the Problem-Dependent Factor (PDF), allows GEMS to take into account how specific an indication is. The PDF varies from near zero for nonspecific indications to one for indications that uniquely identify a single problem. When highly-specific indications are present, GEMS can more precisely diagnose the cause of a problem. The PDF's for a given indication are distributed over the problems it indicates according to how often the indication is likely to be observed when each problem occurs. The Installation Advisor program can be used to specify the PDF value for each combination of an indication and a problem however the default values contained in GEMS were developed and tested as part of the knowledge base development. General guidelines for specifying the PDF are:

PDF	If the indication is present, the problem is ...
1.0	always present
0.8	almost always present
0.6	usually present
0.4	often present
0.2	sometime present
0.0	never present

The contribution of each sensor to the measure of belief in a problem is calculated by multiplying the Problem-Independent Factor by the Problem-Dependent Factor. For example, if a slot temperature reading is an indication of a possible cooling blockage in a particular stator bar with a PDF of 0.6, and the slot temperature has risen to the alarm level (resulting in a PIF value of 0.4), then the measure of belief calculated by GEMS for this problem would be 24% (0.6 X 0.4). The slot temperature sensor deviation could also indicate many other problems to GEMS. Each would have a PIF of 0.4 (the sensor is at the alarm level) and a PDF which would vary with the specificity of this sensor to the particular problem. Thus a number of problems may be diagnosed, each with a different confidence level.

The actual confidence factor generated by GEMS for a particular problem diagnosis is obtained by combining the measures of belief of each abnormal sensor indication using an algorithm similar to that used in Mycin [5]. For example, the confidence factor for a problem with two indications with measures of belief MB1 and MB2 would be calculated as:

$$CF = MB1 + ((1-MB1) * MB2)$$

This normalization algorithm ensures that confidence factors for any given problem never go beyond 100%. In the example above, if a second sensor indication of a blocked cooling problem (for example a high stator winding hose output temperature reading) was present and contributed a measure of belief of 30%, then GEMS confidence in diagnosing a blocked cooling problem would be increased to 47% (0.3 + (1 - 0.3) * 0.24).

GEMS operation is much more complicated than this simple example suggests: GEMS must consider many problems at one time with each having many more than two indications. Sensor ranges for alert, alarm, or limit are calculated in real time, often as a function of other sensor inputs (for example the alert and alarm levels for the stator winding slot temperatures are calculated as a function of the stator current and the bulk coolant inlet temperature). In some cases an aggregate indication may be calculated from multiple sensor readings throughout the generator. The trend of a particular sensor, rather than the absolute range of the sensor may also be of importance. Finally, a problem diagnosed in one subsystem of the generator may be used as an indication for a different problem in another subsystem.

When responding to a particular problem, the turbogenerator operator must consider other factors beyond confidence in his diagnosis of the problem. Both the urgency and severity of the problem play key roles in determining the actions and the speed with which the operator must react. Although GEMS may determine a particular problem is occurring with a very high confidence level, the problem may not be severe in terms of its consequences to the generator, or may be developing slowly and therefore would not require immediate operator action. On the other hand, GEMS may indicate a possible problem to the operator with a very low confidence, however, the consequences to the generator if the problem is actually occurring may be severe. Therefore, an important part of the GEMS diagnosis, is to inform the operator of the severity of any problems detected by GEMS as well as the urgency with which he must react.

The urgency of a problem, in most cases, can be determined by how quickly the particular sensors indicating that problem are changing. If the sensors are changing slowly, the operator may have time to maneuver the unit or take some further diagnostic steps to more closely determine the specific problem occurring. If the sensors are fast approaching their maximum limits, the operator must take immediate corrective action. For each of the problems diagnosed by GEMS, key sensors have been identified to be used to calculate the problem urgency. Urgency for a particular problem is defined as the reciprocal of time to reach limit level for those key sensors identified as critical to that problem. Calculation of the time remaining before a sensor reaches its limit level is based on extrapolation of the recent trend of the indication.

The urgency is then normalized to discrete levels between zero and one (with one indicating a more urgent problem) and displayed to the operator along with GEMS confidence of diagnosis.

URGENCY	TIME REMAINING
1.0	0-3 minutes
0.8	3-10 minutes
0.6	10-20 minutes
0.4	20-60 minutes
0.2	hour

Determining the severity of a particular problem is more difficult than determining urgency. For example, a partially plugged strainer in the stator water cooling system may only become a severe problem when the blockage is large enough to affect cooling to the stator winding (at this point the problem also becomes more urgent since stator winding temperatures would be moving upwards). In effect, severity and urgency are closely related. In the GEMS system, the severity rating of a problem increases with the potential physical damage that could result from ignoring the problem. Problems that are considered more severe are those that could cause more extensive damage to the generator if left uncorrected. Using the installation advisor program, severity has is specified according to the following discrete levels.

Range of Severity:

1.0 - extended generator outage
0.8 - damage to the generator
0.6 - de-rating of the generator
0.4 - partial loss of generator life
0.2 - no adverse effects to the generator

The operator display combines the confidence, urgency, and severity of a problem diagnosed by GEMS with advice and corrective action. Operator advice messages are built from text which can be customized through the use of the Installation Advisor program. For every suspected generator problem, GEMS provides:

- A description of the suspected problem and the confidence in the diagnosis.
- A description of the severity of the problem including the damage that could result if the problem is left uncorrected.
- An indication of the urgency of the problem based on the time before critical sensors reach their maximum limits.
- Recommendations on diagnostic actions that could be taken to further confirm the problem. These recommendations would only be useful if the problem urgency is low giving the operator sufficient time to respond.
- Recommendations for immediate corrective action assuming little time is available for diagnostic actions.

Installation Advisor Program

The Installation Advisor program is used to configure the GEMS monitoring program for a particular generator site. The installation process for GEMS must be undertaken for each new generator site. Information on the type and location of sensors, algorithms for calculating alert, alarm, and limit ranges, operator advice messages, machine design characteristics and modelling information, etc, must all be specified before GEMS can operate correctly. This information is requested through a table driven user interface. Using simple rules, a particular generator configuration is checked for consistency as it is being developed. The Installation Advisor program is also written in the ART expert system shell.

Information obtained through the installation process is organized and stored in three data files for later use by the Monitoring Program; the Sensor Description File which describes the type, location, units, valid operating ranges, graphical plotting ranges, etc, for each sensor; the Generator Description File which contains critical modelling data about the generator (for example the number of parallels in the stator winding or the type of exciter on the unit); and the Utility Policy File which contains specific operator actions and descriptions particular to the utility where GEMS is to be installed. This information is then read by the Monitor Program and used to re-configure the expert system knowledge base. In some cases, whole sections of the knowledge base may be activated or deactivated. For example, if the particular generator being monitored uses a static excitation system then all rules pertaining to rotating exciters would be disabled. As well, the Installation Advisor program is structured in a hierarchical manner so that specific configuration questions relating to, for example, rotating exciters would not be activated once the user specifies a static excitation system is being used.

HARDWARE DESCRIPTION

Because the GEMS software (including the man-machine interface) is being entirely written within the ART expert system shell and Common Lisp, the software can be readily ported to any of a number of Unix workstations. This eliminates the need for a specialized Lisp machine and allows GEMS to be economically delivered as an in-plant monitoring system. By dividing the monitoring program into two separate parts and using the control structure described above, GEMS will not have to run on an expensive mainframe computer in order to update its advice to the operator at three-minute intervals, but will be able to achieve this speed when running on a relatively inexpensive workstation. With current workstation memory size and processing capabilities, one monitoring computer is required for each generator to be monitored by GEMS.

Data acquisition for GEMS can be accomplished by one of two means. In older plants, where a great deal of the generator sensor data may not be available in digital form, a dedicated acquisition system is necessary. A process in the GEMS monitoring computer is then used to communicate with this acquisition system and obtain sensor snapshots. The Installation Advisor program is customized to handle a specific data scanner (Fluke Helios I) and will set up the necessary configuration files and sensor conversion algorithms to be downloaded to this device. In plants where the generator sensors are already available and converted to engineering units by a plant computer, a data link can be established between this computer and the GEMS monitoring computer. If a process can be written for the plant computer to allow it to emulate the Fluke data logger, then no changes are necessary to the GEMS code. If this is not possible, some customization of the GEMS data acquisition program would be necessary.

Regardless of which acquisition technique is used, the interface between the GEMS monitoring system and the generator sensors is handled through a standardized file format. Data snapshots are queued in this file system for processing by the monitoring program. This architecture allows for easy testing of GEMS in an off line manner. An independent program called the Generator Input Simulator Program (GISP) has been written and can be used to create test scenarios. These test scenarios consist of a time series of data snapshot files with abnormal sensor indications generated in them. A graphical interface is used by the GISP to plot and modify sensor indications using a pointing device (mouse). This simplifies the examination and creation of multiple tests cases using the GISP.

CURRENT STATUS

A prototype GEMS is now under construction. To simplify and modularize the software and knowledge engineering tasks, generators have been divided into a number of subsystems. Knowledge engineering has been completed for the stator winding, excitation, rotor, and core subsystems. The overall framework for all of the programs described above has been completed and rules encompassing the stator winding subsystem have been written. Preliminary testing of this software has begun using the GISP. Software development is done on a Symbolics Lisp machine and ported for delivery on a Sun 3/60 workstation. Two installations of the prototype system are planned. The first installation of GEMS will be made on a 500 MW turbogenerator at the Nanticoke Thermal Generating Station of Ontario Hydro (Canada) in May of 1989. A second installation is planned for a 850 MW turbogenerator at the Oswego plant of Niagara Mohawk Power Corporation (USA) early in 1990.

CONCLUSION

This paper describes the design of a real time expert system for monitoring of turbogenerators. Many of the techniques employed in this application could be extended for use in other monitoring applications. Although the basic feasibility of an expert system monitor for turbogenerators is obvious, GEMS presents many technical challenges associated with real time processing capabilities and the need for an adaptive system which can be applied to generators of varying design, manufacture, and operating environment. The successful deployment of this system will clearly demonstrate the capability of applying expert systems to monitoring and diagnostic applications in the power industry.

REFERENCES

1. EPRI EL3564-SR, "Workshop Proceedings, Generator Monitoring and Surveillance", August 1984.

2. A.J. Gonzalez, R.L. Osborne, C.T. Kemper, S. Lowenfeld, "On-Line Diagnosis of Turbine-Generators Using Artificial Intelligence", IEEE/PES for 1985 Winter Meeting.
3. A. Clapis, M. Gallanti, A. Stefanini, "Expert Systems in Plant Diagnostics, A Practical Application", Automazione E Strumentazione, March 1984.

4. E. Taymans, F. Bastenaire, "Operator Advisor, An object oriented expert system for process control", AIM Conference Proceedings, Power Stations, pg 77.1, 1985.

5. E.H. Shortcliffe, "Computer Based Medical Consultations, MYCIN", American Elsevier Publishing Company Inc, 1976, pg 159.

Cooperating Expert Systems for Diagnoses of Electrical Apparatus

MIGUEL A. MARIN
Institut de Recherche d'Hydro-Quebec
1800, montee Ste-Julie
Verennes, Quebec, Canada J0L 2P0

JEAN-LOUIS JASMIN
Essais et Expertises Techniques
Hydro-Quebec
5655, rue de Marseille
Montreal, Quebec, Canada H1H 1J4

ABSTRACT

This paper presents a prototype expert system SEDA-TRANSFO, which implements the cognitive cycle followed in the maintenance and troubleshooting of a high-voltage transformer. It comprises five cooperating modules, i.e. five individual rule-based expert systems for operation, inspection, dissolved-gas analysis, tests and repairs, and a sixth module, analyses, which uses diagnoses emerging from the five modules in order to issue a verdict.[1] The concept of cooperating expert systems is particularly useful in this context.

The first five modules of SEDA-TRANSFO are already operational while analyses (also ruled-based) is under development. The shell used is Rulemaster-2 (Radian Corporation, Austin, Texas). Modules 1 to 5 and a functional description of module 6 are undergoing field tests in various regions of Hydro-Québec to complete the information needed to develop the final product.

1. INTRODUCTION

To increase the availability and life span of its electrical apparatus, Hydro-Québec follows a diagnostic process which indicates the status of the apparatus in question and any maintenance or troubleshooting activities to be undertaken. This process may be viewed as a cognitive cycle involving the following steps: 1) work requisition, 2) knowledge of the status of the apparatus in question, 3) validation of the status by physical inspection, 4) tests to confirm deterioration and/or previous diagnoses, 5) working plan of the activities to be performed, 6) execution of the working plan, and 7) updating of the maintenance program and/or determination of the events that result in the need for a work requisition. It is interesting to note that this cycle is independent of the apparatus concerned and that it produces a diagnosis and associated activities at each step. Using these diagnoses, the maintenance personnel should then be able to identify the cause of the malfunction and assess the urgency of the intervention.

Each step of this cognitive cycle can be implemented as a rule-based expert system, each producing a diagnosis with an associated activity. These expert systems may be used independently at the user's request (e.g. as aids) at any time but they also produce the information needed by another expert system (likewise rule-based) called ANALYSES, whose mission is to issue a verdict[1] on the status of the apparatus in question. In this context, the concept of cooperating expert systems is obviously useful.

---

[1] The concept of verdict in cooperating expert systems, which is related to the structure and implementation of analyses is discussed in more detail in [1].

This paper presents a prototype expert system SEDA-TRANSFO, which implements the concepts outlined above for a high-voltage transformer. SEDA-TRANSFO comprises five cooperating modules, i.e. five individual expert systems: operation, inspection, dissolved-gas analysis, tests and repairs, and a sixth module, analyses, which uses the diagnoses emerging from the five other modules in order to issue a verdict.

The following Section 2 presents the concept, motivation and scope from the domain viewpoint. The concept of maintenance and troubleshooting by diagnoses, on which the SEDA-TRANSFO architecture is based, meets a specific need, namely unification of the different aspects covered by the different expert systems, and the needs of the maintenance and troubleshooting functions. It is motivated by the Apparatus Department's awareness that traditional methods have their limitations.

Section 3 is concerned with SEDA-TRANSFO itself. First it is situated within a general architecture, called SEDA, which integrates a family of expert systems and existing corporate and local databases for the function of the apparatus. Then the architecture of SEDA-TRANSFO and its components are presented. Typical results are chown in Section 4. Case 2, namely, gas relay tripping + differential relay tripping + gas alarm, is presented and discussed. Section 5 presents aspects of the implementation of SEDA-TRANSFO with Rulemaster-2 and the experience obtained. The conclusions, Section 6, summarize the experience gained with such a prototype and describes directions for future development.

2. MAINTENANCE AND TROUBLESHOOTING BY DIAGNOSIS

The concept used in the design and implementation of SEDA-TRANSFO is based on the principle of maintenance and troubleshooting by diagnosis [2]. Contrary to prescheduled maintenance intervention, the principle of maintenance by diagnosis is defined as intervention depending on the state of the apparatus and its past history, from which a diagnosis and associated action may be derived. Troubleshooting by diagnosis is similarly defined: based on the status of the apparatus and other facts at the moment of failure, a diagnosis and action is deduced.

Maintenance personnel apply this principle to troubleshooting activities by following a cognitive cyclic process, which may be visualized as shown in Figure 1. The process, which is independent of the type of apparatus, starts with a work requisition, followed by acquisition of the knowledge regarding the status of the apparatus. The knowledge is then validated by a physical inspection. At this stage, it is sometimes possible already to conclude on a diagnosis and action without completing the cycle. Other times, the diagnosis is preliminary and needs to be confirmed by specific tests on the apparatus. Next in the cycle is the working plan (actions). The maintenance program may be affected by the execution of these actions and is therefore amended so as to produce corresponding triggering events which will produce the required work requisitions in future. The cycle is thus completed. The central circle in Figure 1 represents the maintenance personnel's analysis of situations, based on experience, as they execute the cycle.

The fact that this cognitive cyclic process is based primarily on the experience of maintenance personnel and is applicable to any type of apparatus gave rise to a pilot project with a twofold objective. The first was to prove that expert-systems technology can be applied advantageously to maintenance and troubleshooting functions. The second was to propose a general development concept for the implementation of an entire family of expert systems, at the level of an installation (say, a substation), covering five types of apparatus, namely, transformers, circuit breakers, rotating machines, and low-voltage and high-voltage auxiliary equipment.

For the purpose of prototype implementation, a power transformer undergoing troubleshooting after automatic tripping was chosen. In this situation, only the following four typical cases were to be considered: case 1: tripping by gas relay; case 2: tripping by gas relay + differential relay + gas alarm; case 3: tripping by differential relay, and case 4: tripping by overload + gas alarm. The prototype was to be flexible enough to behave merely as an aid to the user and was never to replace the latter's decision-making. Also, if possible, access to corporate databases for equipment data was to be provided in order to benefit from corporate data processing facilities but this objective was soon abandoned when it was realized that the data needed to feed the expert systems were resident in incompatible systems. The design of suitable interfaces was beyond the scope of the pilot project. However, as shown in the next section, a convenient and flexible data acquisition facility, i.e. printable questionnaires, was provided and a general architecture was proposed for this purpose.

3. SEDA-TRANSFO

3.1 Definition

The prototype SEDA-TRANSFO is a rule-based expert system of the demonstrator type which implements the troubleshooring process used by maintenance personnel in a power transformer automatic-tripping situation. It is part of a global architecture, called SEDA (Système expert de diagnostics d'appareillage), whose objective is to provide an approach and a concept for implementing a set of cooperating expert systems relative to the electrical apparatus of an installation, e.g. a substation (see Figure 2). Thus, SEDA is composed of two major parts: 1) SEDA-G, which acts as the front-end and interfaces with the different corporate databases, and 2) the SEDA-PX, SEDA-PY, SEDA-PZ expert systems corresponding to apparatus PX, PY, PZ. These expert systems are both independent and cooperating at the same time. Each SEDA-PX contains four expert subsystems covering the four different aspects of the apparatus: electrical, mechanical, civil and transportation, which are not necessarily related but, in certain cases, may have a strong link and therefore cooperate.

The prototype SEDA-TRANSFO is the first SEDA-PX developed within the framework of the SEDA architecture. The hashed area in Figure 2 represents the part corresponding to the present version of this prototype.

3.2 Architecture

The architecture of SEDA-TRANSFO was inspired by the practical cyclic process used by maintenance personnel in troubleshooting, as discussed in Section 2. The different levels of expertise involved in the execution of this cycle call for a very flexible and friendly disign to allow either independent or sequential use of the modules. Thus, the architecture (Figure 3) comprises six modules, each independent of the others, which can be called via a main menu. Their nature and selection result naturally from the cycle process shown in Figure 1: module 1: Operations; module 2: Inspection; module 3: Tests: Insulation fluids, dissolved gases; module 4: Tests: Equipment; module 5: Reconditioning; module 6: Analyses. Module 0 (not listed above) contains a general description of the prototype, its functions and its limitations.

Modules 1 to 5 have a similar structure. Each contains four parts: a description of the approach taken by the module in question; a questionaire, which can be printed, to help the user gather the required input data; a set of questions-answers displayed on the screen as the user enters the requested input data; and a summary of entered data, diagnoses and corresponding actions, which may also be printed.

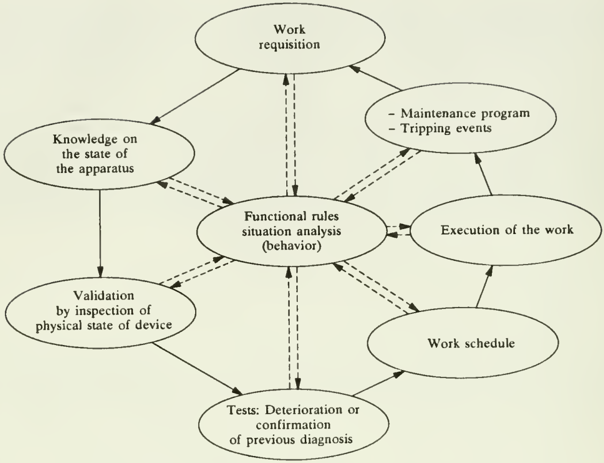


Figure 1. Cycle of Maintenance and Troubleshooting Actions by Centralized Diagnostics.

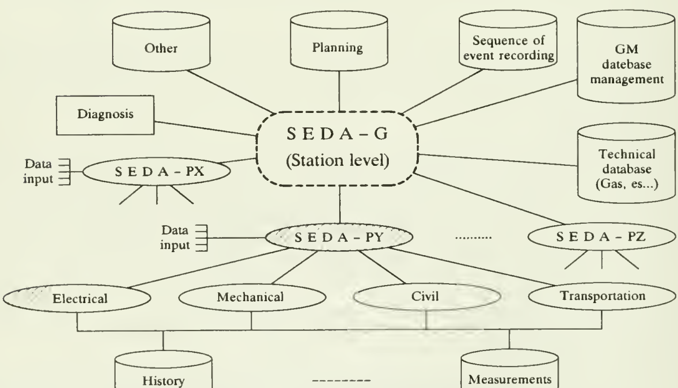


Figure 2. Expert Systems for Diagnosis of Equipement SEDA ARCHITECTURE.

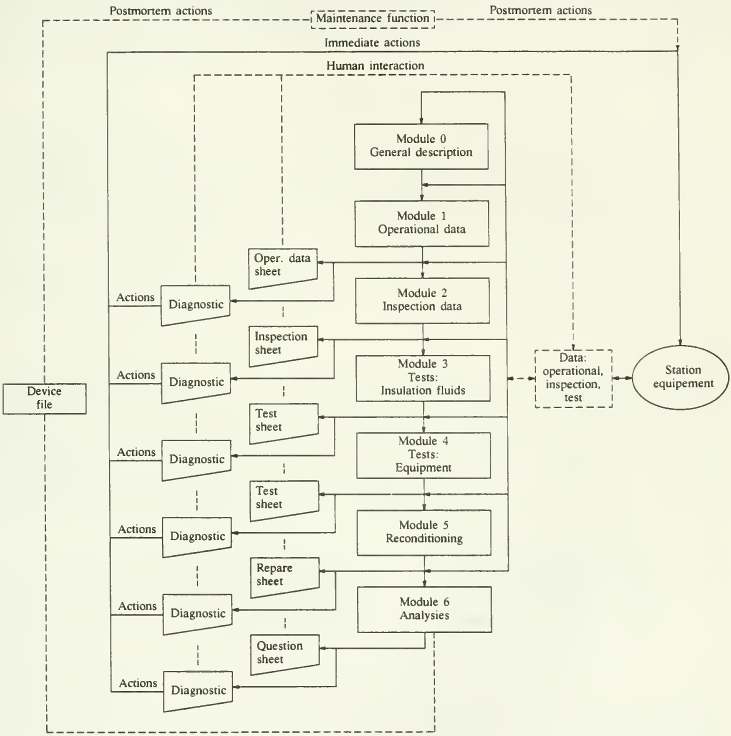


Figure 3. SEDA-TRANSFO Architecture and its Context.

Module 6 differs from the others in that its vocation is to assist the user in the analysis of the cause of the automatic tripping and/or failure. The present version of this module simply provides a hint to the user on how to pursue the analysis should the diagnoses given by the remaining modules not be conclusive. The present implementation, discussed more in detail in [1], produces a verdict on the apparatus in question using the diagnoses of the other modules, and displays previous cases ("jurisprudence") upon request for perusal by the user who issues the final verdict. In its ultimate version, this module will contain an aging model of the apparatus and should have access to corporate databases.

3.3 Components

Modules 1 to 5 will now be discussed in greater detail.

The mission of module 1, Operations, is to determine the operationing data relative to the apparatus in question together with its state after automatic tripping has occurred. This is accomplished by asking the following types of question on the screen: identification of the apparatus, its location and type of intervention (protection zone displayed as a memory aid to the user); type of protection tripping; type of alarm; type of reading, e.g. overload, overvoltage, ground current; type of observation noted, e.g. explosion, fluid overflow, injured person. At the end of this questionaire, a set of corresponding heuristic rules is executed, which produces on the screen a summary of the entered data and the associated actions or advice to be taken by the operator. These two outputs can be printed by activating the PRINT-SCREEN key.

Module 2 covers two types of physical inspection; Inspection A covers seven transformer items, i.e. oil level, overpressure devices, main tank, control box, bushing gas relay and dryer, while Inspection B is concerned with the protection zone, i.e. circuit breakers, lightning arresters and switches. The module is executed in two parts: if inspection A is normal, then the computer bypasses it and displays the inspection B questionnaire. It terminates with a summary of answers to questions and a list of corresponding actions/recommendations. It is interesting to note that after a question is answered the system responds with advice. The user may at this time opt to continue or to abort, depending on his or her objective and knowledge of the situation.

Module 3, Test; Insulation Fluids, is designed to include several types of such tests as they become available. A menu is therefore provided for this selection when called but, for the time being, only dissolved-gas analysis has been implemented. Two methods are used: Duval's method [3,4] and the IEC (International Electrotechnical Commission) method [5]. After entering the gas concentration, a diagnosis is given, together with a summary of input data for the two methods. Experimentation with laboratory test data revealed Duval's method has a broader coverage of cases than the IEC method. The second part of this module is converned with the severity or potential danger of a transformer fault as a function of the dissolved-gas concentration and the age of the transformer [4]. This conclusion is based on empirical data and heuristic rules currently under-going field tests.

Equipment tests are covered in Module 4 and comprise four types: DC insulation, TTR (transformer turns ratio), AC insulation/magnetization current and DC resistance. The module is organized in such a way as to allow the user to call any desired test as many times as needed. Depending on the answers to the questions related to the readings and conditions experienced during the execution of the test, advice is given to the user for immediate action, if desired. As with the other modules, a summary is presented after each test called. In this module, a special effort was made to provide the user with as much useful information as special effort was made to provide the user with as much useful information as possible (not readily available) performing or interpreting the tests. For example, in the case of transformer drainage, a prompt is displayed with a list of operator safety measures.

Finally, module 5 is concerned with a particular working method to perform internal inspections of a transformer. First, the method of drining the transformer is given, together with safety measures. Then the inspection procedure, based on experience and standards, is given. It comprises eight items: main tank, off-load tap switch, on-load tap switch, windings, current transformer, terminals, magnetic circuits and surge arresters.

4. TYPICAL RESULTS

Typical results obtained with SEDA-TRANSFO are illustrated in Figures 4 to 7. The case studied is case 2, namely, gas relay tripping + differential relay + gas alarm, where a transformer is supposed to be in a situation such that gas tripping, differential protection and gas alarm were all detected. Since SEDA-TRANFO is an off-line, stand-alone system, it can be interrogated at any time after the fact.

Figure 4 shows the results of module 1. Note that they indicate the occurrence of an explosion, oil spill and injuries. Therefore, the diagnosis calls for actions involving the utility's Apparatus, Safety and Environment departments.

Figure 5 presents the results of a physical inspection and the associated recommended actions. Note that some actions give an immediate intervention plus a next step. For example, ACTION A.5-a, which occurs when the oil level in the bushings is low, recommends that oil be topped up in the bushing, that an insulation test be performed and that module 4 be used to interpret the test results. In this way, the different modules guide or cooperate with the user step by step. ACTION A.6 recommends dissolved-gas analysis. According to module 3 (Figure 6), Duval's method indicates high-energy arcing but, since the age of the transformer is 15 years, it is concluded that the fault is not dangerous.

Finally, module 4 (Figure 7) gives the diagnoses and actions associated with the four tests performed. Note that in some cases, such as in Test 2 (TTR), advice and reference to the maintenance manual, i.e. section 7/appendix 5, are given. This manual (text and drawings) can easily be incorporated into the module and prompted upon request.

5. COMMENTS ON RULEMASTER-2 IMPLEMENTATION

RuleMaster-2 [6] is a software tool for building rule-based expert systems which has been developed by Radian Corporation of Austin, Texas. Two features are especially attractive for the diagnostic application in question: the automatic rule generator and Radial, the structured rule language. The rules developed with RuleMaster contain rules induced from examples and/or written directly in Radial. As examples of these features, Figure 8 shows Duval's triangle and Figure 9 its implementation using conditional rule states; Figure 10 represents the implementation of the IEC method using rules induces by examples.

Besides the rule generation facilities, RuleMaster generates code in C-language and produces executable code under MS-DOS, which is deployed on personal computers.

The following information is entered:

- -

INSTALLATION IDENTIFICATION: XXXX XXXXXX
APPARATUS IDENTIFICATION: XXXX XXXXXX
TYPE OF INTERVENTION: 02. - unpredictable fault

- -

Is this information correct? [yes, no] yes

************* **SUMMARY OF OPERATIONS DATA** ************

NOTE: This summary table shows the entered operations data and is related to the following DIAGNOSES and ACTIONS.

TRIP Tx...63* GAS:	YES
TRIP Tx...87* DIFF:	YES
ABNORMAL INDICATION Tx...63 GAS:	YES
EXPLOSION	
OIL SPILL	
INJURIES	

Phase noted: A
GROUND CURRENT amplitude: 999

***** **DIAGNOSES and ACTIONS related to operations data** *****

DIAGNOSIS:

Major fault on phase A
Case 2: Tripping by gas relay + Differential relay + Gas alarm
ACTION: Notify and wait for instructions from: APPARATUS DEPARTMENT
Person in charge: XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
Telephone: XXX-XXXX

ACTION: Notify and wait for instructions from: SAFETY and ENVIRONMENT DEPARTMENTS
Person in charge: XXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
Telephone: XXX-XXXX

************* **END OF DIAGNOSES FROM OPERATIONS** ************

To exit and return to MAIN MENU, press ENTER.

Figure 4. Output of Module 1: Operations, for CASE 2

```
********    SUMMARY OF DIAGNOSES FROM INSPECTION DATA    ********

     - - - - - - - - - - - - - - - - - - - - - - - - - - - -

DATA ENTERED:
    A.1)    Oil level:
    A.1-a)  Expansion tank:     low-level alarm:  low ambient temp.
    A.1-a)  Expansion tank:     low-level alarm:  leak
    A.3-a)  Main tank:          deformed
    A.3-b)  Main tank:          faded
    A.4-b)  Control box:        current-transformer wiring heated
    A.5-a)  Bushings:           oil level low
    A.5-b)  Bushings:           by-pass:   yes
    A.5-c)  Bushings:           discolored: yes
    A.6)    Gas relay:          operation:  yes;  gas:  yes
    B.1)    Circuit breaker(s): break  and/or  by-pass:  yes
    B.4)    Switches:           break  and/or  by-pass:  yes


The ACTIONS corresponding to these data are found on the next page.


To continue press ENTER.

******    ACTIONS RESULTING FROM INSPECTION DATA ENTERED    ******


ACTION:
    A.1)    Fill  (Expansion tank)
    A.1-a)  Locate, repair, fill  (Expansion tank)
    A.3-a)  Main tank deformed  →  MODULE 4
    A.3-b)  Main tank faded  →  MODULE 4
    A.4-b)  Verify wiring continuity of current transformer
    A.5-a)  Fill with oil (bushings).  Insulation test → MODULE 4
    A.5-b)  Insulation test (bushings)  →  MODULE 4
    A.5-c)  Insulation test (bushings)  →  MODULE 4
    A.6)    Gas relay:  operation  →  MODULE 4,  Tests +
            dissolved-gas sample  →  MODULE 3
    B.1)    Circuit breaker(s):  repair
    B.4)    Switches:  repair


*****        END OF DIAGNOSES FROM INSPECTION DATA        ******


To exit and return to MAIN MENU, press ENTER.
```

Figure 5. Output from Module 2: Inspection, for CASE 2.

******** **SUMMARY OF DISSOLVED GAS ANALYSIS** ********

Gases			(ppm)
HYDROGEN	(H2)	:	1
OXYGEN	(O2)	:	24.3200
NITROGEN	(N2)	:	75.5500
CARBON MONOXIDE	(CO)	:	1
METHANE	(CH4)	:	1
CARBON DIOXIDE	(CO2)	:	32
ETHYLENE	(C2H4)	:	1
ETHANE	(C2H6)	:	1
ACETYLENE	(C2H2)	:	1

DUVAL method: ZONE 1: High-energy arcing
IEC method: FAULT NOT DEFINED BY IEC METHOD
CO2/CO: NO PAPER INVOLVED

*** **RECOMMENDATIONS ON POTENTIAL DANGER OF FAULT TO APPARATUS** ***

REMARK: These recommendations are now under study and must be validated. However, they provide an indication of the potential danger of the fault mentioned above, for the apparatus concerned.

To continue, press ENTER.

Does the transformer in question have a tap changer connected to the main tank [yes, no] yes

ADVICE: Recommendation not available
(RETURN continues)

Was the oil sample taken at the bottom of the tank? [yes, no] no

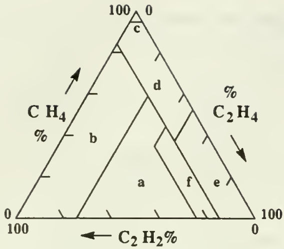
ADVICE: Recommendation not available
(RETURN continues)

What is the age of the transformer, in years? 15
ADVICE: The fault is an arc (ZONE 1 or 2, Duval)
The fault is not dangerous for the apparatus.

Figure 6. Output of Module 3: Tests: Insulation Fluids, Dissolved-gases analysis, for CASE 2.

```
******        SUMMARY OF RESULTS OF DC INSULATION TEST        ******
                              (MEGGER)

STEP 1A - INSULATION TEST:          Reading:  INFINITY
     ACTION:   Continue test  →  STEP 2
    ________________________________________________________________

STEP 1B - CONTINUITY TEST:
          Result:   CONTINUITY
          ACTION:   Continue DC resistance test  →  STEP 4
    ________________________________________________________________

To continue, press ENTER

****   SUMMARY OF RESULTS OF TRANSFORMER TURNS RATIO TEST   *****
                               (TTR)

STEP 2 -  Result:  RATIO DIFFERENCE - BETWEEN PHASES
          CAUSE may be:   a)  off-load tap changer
                          b)  on-load tap changer
          ACTION:  Verify mechanism and proceed with resistance
                   test (section 7/appendix 5).
    ________________________________________________________________

To continue, press ENTER.

******        SUMMARY OF RESULTS OF AC INSULATION TEST        ******
                              (DOBBLE)

STEP 3A - Result:  Reading UNSTABLE BETWEEN WINDINGS
          CAUSE:   Short-circuit possibility
          ACTION:  Confirm with DC resistance test  →  STEP 4
    ________________________________________________________________

To continue, press ENTER.

*****   SUMMARY OF RESULTS OF MAGNETIZATION CURRENT TEST   *****

STEP 3B - Result:  Reading IMPORTANT VARIATION BETWEEN PHASES
          CAUSE:   Possible partial short-circuit in windings
    ________________________________________________________________

To continue, press ENTER.

*****        SUMMARY OF RESULTS OF DC RESISTANCE TEST        *****
                        (RESISTANCE BRIDGE)

STEP 4 -  Result:  Reading LARGER
          CAUSE:   1) Possibility of partially open windings
                   2) Loose connection on the taps
                   3) Loose connection on the joints inside the main tank
                   4) Loose connection on the external connections

          ACTION:  Continue performing more precise tests on each
                   element of these sets.
    ________________________________________________________________

To continue, press ENTER.
```

Figure 7. Output of Module 4: Tests: Equipment, for CASE 2.



a.	High-energy arcing ($I > 20\ I_n$)
b.	Low-energy arcing, tracking
c.	Corona discharges
d.	Hot spots, T < 200°C
e.	Hot spots, 200 < T < 400°C
f.	Hot spots, T > 400°C

Triangle coordinates:

$\% \, C_2H_2 = \frac{100x}{x+y+z}$; $\% \, C_2H_4 = \frac{100y}{x+y+z}$; $\% CH_4 = \frac{100z}{x+y+z}$

with $x = [C_2H_2]$; $y = [C_2H_4]$; $z = [CH_4]$ in ppm

Figure 8. Duval's Triangle, Calculations and Interpretation of Zones [3].

```
    STATE: duval
    IF p_C2H2 < 10 IS
"T": (null, zone_345)
ELSE (null, zone_126)

    STATE: zone_345
    IF ((( p_CH4 > 95 ) and ( p_C2H4 < 5 )) and (p_C2H2 < 5 )) IS
"T": ("ZONE 3 : Decharges couronnes" -> zone; 3 -> z; prints "\n"; prints zone,CEI)
ELSE IF (( p_C2H4 > 50 ) and ( p_CH4 < 50 )) IS
"T": ("ZONE 5: Points chauds  200 < T < 400 C" -> zone; 5 -> z; prints "\n"; prints zone, CEI)
ELSE ("ZONE 4 : Points chauds < 200 C" -> zone;  4 -> z; prints "\n"; prints zone, CEI)

    STATE: zone_126
    IF (( p_CH4 < 85 ) and ( p_C2H4 < 25 )) IS
"T": ("ZONE 2 : Arcs de faibles energie" -> zone; 2 -> z; prints "\n";prints zone, CEI)
ELSE IF ((( p_CH4 < 45 ) and ( p_C2H2 < 25 )) and ( p_C2H4 > 40 )) IS
"T": ("ZONE 6 : Points chauds > 400 C" -> zone; 6 -> z; prints "\n"; prints zone, CEI)
ELSE ("ZONE 1 : Arcs de forte energie" -> zone; 1 -> z; prints "\n";prints zone, CEI)
```

Figure 9. Implementation of Duval's Triangle with RuleMaster-2.

```
STATE: CEI_1
   ACTIONS:
      0              ["CAS 0:  PAS DE DEFAUT" -> region]
      1              ["CAS 1:  DECHARGES PARTIELLES DE FAIBLE DENSITE D'ENERGIE" -> region]
      2              ["CAS 2:  DECHARGES PARTIELLES DE FORTE DENSITE D'ENERGIE" -> region]
      3              ["CAS 3:  ARCS DE FAIBLE ENERGIE" -> region]
      4              ["CAS 4:  ARCS DE FORTE ENERGIE" -> region]
      5              ["CAS 5:  POINT CHAUD   < 150 degres C" -> region]
      6              ["CAS 6:  POINT CHAUD  150 < T < 300 degres C" -> region]
      7              ["CAS 7:  POINT CHAUD  300 < T < 700 degres C" -> region]
      8              ["CAS 8:  POINT CHAUD  T > 700 degres C" -> region]
      otherwise      [("\BDEFAUT NON DEFINI PAR LA METHODE CEI\N" -> region,paper)]

   CONDITIONS:
      r1             [ace_eti]    { 0  1  2 }
      r2             [met_hid]    { 0  1  2 }
      r3             [eti_eta]    { 0  1  2 }

   EXAMPLES:
      0              0            0              => (0,paper)
      0              1            0              => (1,paper)
      1              1            0              => (2,paper)
      1              0            1              => (3,paper)
      2              0            1              => (3,paper)
      2              0            2              => (3,paper)
      1              0            2              => (4,paper)
      0              0            1              => (5,paper)
      0              2            0              => (6,paper)
      0              2            1              => (7,paper)
      0              2            2              => (8,paper)
```

Figure 10. Implementation of the IEC Method [5] with RuleMaster-2.

The reasons for selecting RuleMaster for this application were:

1) Rules induced by examples and (un)conditional rule states provide very easy way to implement rule-based, forward-chaining, inferencing of the type used in diagnoses of equipment, where a set of facts is associated with a diagnosis and a specific action or recommendation.
2) SEDA-TRANSFO and all its derivatives had to be deployed in an environment where PCs are already in use for other applications, such as accessing corporate databases. MS-DOS and PCs were therefore fixed requirements from the beginning.
3) The architecture of SEDA (Figure 2) calls for access to corporate databases and, eventually, to special person/machine interfaces. Since RuleMaster gerates C-language code, the use of specially programmed features in C could be easily linked with the expert systems SEDA-PX, as they become available.
4) The fact that there is a system-call utility in RuleMaster allows it to call MS-DOS functions, such as type.../MORE directly, which proved very helpful in displaying large quantities of text on the screen.
5) The explanation facility was not a great concern for this level of development because the end-user did not require explanations and was satisfied with summaries (data entry/diagnosis/action). Actually, the explanation facility was turned off before delivery, although during the development stage it was used extensively. The rule inconsistency warning, expecially regarding the examples used to generate rules, and the tracing facility were very helpful.
6) RuleMaster was known from previous applications, so that it was easy to rapidly implement, the knowledge of the experts, as it became available, for verification purposes.
7) The interfacing facility of RuleMaster-2 was not used because of the special requirements of the application, one of them being the use of the French language. The inability to incorporate French punctuation was, and still is, of concern to the developers. This is a minor problem, however, which can be easily overcome.
8) The hardware/software investment needed to begin developing SEDA-TRANSFO was very low, since all that it required was to purchase RuleMaster-2 under MS-DOS. The PCs were already available at all potential user sites.

6. CONCLUSION

The prototype expert system SEDA-TRANSFO presented in this paper implements the cognitive cycle followed in the maintenance and troubleshooting activities for a high-voltage transformer. This cycle is the same for all types of electrical apparatus and thus provides a general concept on which to base the development of a whole family of expert systems, SEDA-PX..., SEDA-PZ, covering transformers, circuit breakers, rotating machines, HV and LV auxiliary equipment.

SEDA-TRANSFO is described as being part of a general architecture, called SEDA, whose objective is to provide a concept for implementing a set of cooperating expert systems covering all the electrical apparatus of a given installation. This architecture contains a front-end, SEDA-G, whose vocation is to interface with corporate databases and to format the data required for the differnt SEDA-PXs.

The architecture of SEDA-TRANSFO comprises six modules, all rule-based expert systems in themselves, which may be accessed at any time by users via a main menu depending on needs and on the knowledge that they may have of the situation under study. In this sense, these modules cooperate in achieving the ultimate goal, final decision or verdict [1] to be taken about the apparatus in question. Module 6, analyses, is responsible for providing the user with this verdict based on the diagnoses emerging from each of the other modules.

A typical output for case 2: gas relay tripping + differential relay + gas alarm, of a transformer was given as an illustration of the capabilities of SEDA-TRANSFO at this stage of development.

RuleMaster from Radian Corporation (Austin, Texas), was used in the implementation of SEDA-TRANSFO and some comments were given on the authors' experience gained with such a development tool.

A copy of SEDA-TRANSFO is now deployed in each administrative region of Hydro-Québec. Comments received so far are very encouraging. They refer primarily to the availability in one place (the screen) of very useful and much needed information for deciding what to do with a particular item of apparatus under certain conditions. This prototype also provided an opportunity to prove the feasibility of the domain concept and the software architecture.

Finally, a major effort is now underway to finalize SEDA-TRANFO, continue the development of the SEDA-PXs and start work on SEDA-G.

ACKNOWLEDGMENTS

The authors wish to acknowledge the contribution of domain experts G. Bourguignon, G. Cantin, M. Duval, R. Sirois and G. Poisson, the project leader A. Jolicoeur and, finally, the authors' respective department heads, R. Blais and J.-R. Valotaire, for their guidance, encouragement, continuous support for the project, and comments on this paper.

REFERENCES

[1] M.A. Marin and J.-L. Jasmin. " The Concept of Verdict in Cooperating Expert Systems for Diagnoses of Electrical Apparatus," Conference on High Technology in the Power Industry, IASTED, Valencia, Spain, July 4-7, 1989.

[2] J.-L. Jasmin and J-P Henry. "Entretien prévionnel: Philosophie et projet-pilote du service Appareillage, Région Maisonneuve," Hydro-Québec, April 1984.

[3] M. Duval, "Fault Gases Formed in Oil-Filled Breathing EHV Power Transformers," IEEE-PES Conference Paper C 74 476-8, 1974.

[4] M. Duval et al., "Acceptable Dissolved Gas-in-Oil Concentration Levels in Power and Instrument Transformers versus Age," Minutes of the 1989 Doble Conference. April 1989.

[5] International Electrotechnical Commission. "Interpretration of the Analysis of Gases in Transformers and other Oil-filled Electrical Equipment in Service," IEC Publication 599, 1978.

[6] RuleMaster-2. Software Tool for Building Expert Systems, Radian Corporation, Austin, Texas, September 1987.

[7] A. Gupta and B. E. Prasad. Principles of Expert Systems. IEEE Order Number: PC 0228-7, IEEE Press, 1988.

[8] A. Gupta and B. E. Prasad. Microcomputer-based Expert Systems. IEEE Order Number: PC0231-1, IEEE Press, 1988.

Expert System for On-Line Monitoring of Large Power Transformers

THOMAS H. CROWLEY, WAYNE H. HAGMAN, RICHARD D. TABORS, and CHATHAN M. COOKE
Laboratory for Electromagnetic and Electronic Systems
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139, USA

ABSTRACT

The failure of large power transformers is an area of significant cost and concern for electric utilities. Often transformer failure is catastrophic, because there is no early warning of incipient failures. This paper first discusses the economic value of a transformer monitoring system and then presents a concept for an on-line transformer performance monitoring system with dramatically increased sensitivity over conventional threshold methods for the detection and diagnosis of incipient failures. The concept centers on continuous on-line monitoring of several subsystems in a transformer. Anomalies in subsystems are detected by comparing the actual operation with adaptive models of what is normal for the transformer. Detection and diagnosis of incipient failures is performed by cross-correlating anomalies and other information about subsystems, then matching the results to failure modes using an expert system approach. Research on the detection portion of the system is essentially complete; however, the diagnosis portion involving the expert system is the subject of ongoing work. A prototype laboratory implementation of the on-line detection portion of the system is described; the implementation is designed around two 80286-based personal computers and the UNIX operating system. Results of on-line tests, monitoring a 50 kVA transformer in the laboratory, and indicating increased sensitivity to an incipient failure, are presented.

INTRODUCTION

The failure of large power transformers is an area of significant concern for electric utilities. Transformers are major elements in power generation and transmission systems. Failures, particularly those which come without warning, cause service disruptions which are frequently difficult to circumvent and may cost millions of dollars in replacement fuels or customer outages. The present failure rate of large transformers in the U.S. is about 2% per year [1]. However, the tremendous cost of failures, even at such a low rate, causes many utilities to purchase spare transformers and install redundant equipment; tying up capital and manpower needed elsewhere.

The ability to foresee, or at least identify the existence of, incipient transformer failures before they become catastrophic is highly desirable. The benefits of such early warning fall broadly into four categories:

- Prevention of catastrophic failures and sudden outages
- Optimization (and cost minimization) of maintenance
- Estimation of remaining life
- Better utilization of capacity

A large electrical transformer is a complicated mechanism, the condition of whose constituent parts cannot be readily evaluated, if at all, from external observation. The identification of incipient failures must therefore be achieved through the monitoring of internal characteristics. Past experience, however, has illuminated the complexity of the coupling between failure processes and subsystem (windings, insulation, oil, core, sensors, etc.) responses, or signatures. Even though the internal environment and external operating conditions of a large power transformer make data acquisition and analysis extremely difficult tasks, accurate performance monitoring of the internal condition of an in-service transformer remains nonetheless attractive.

Under Electric Utility Sponsorship[1], the Laboratory for Electromagnetic and Electronic Systems at MIT has undertaken a research program with the broad goal of establishing advanced technologies to significantly improve the reliable monitoring of large in-service power transformers, allowing for the detection of incipient failure conditions. This effort can be viewed in terms of four areas:

- Development of Basic Sensors and Understanding of Sensor Signals
- Understanding and Modeling the Operation of Transformer Subsystems
- Development of Integrated Monitoring System Software and Hardware
- Testing of Sensors and System on a 50 kVA Transformer

An adequate description of the work carried out in and amongst these four areas would fill a small book; this paper deals with the results of a portion of the work listed above, specifically: Development and Testing of Integrated Monitoring System Software and Hardware[2].

Accurate, in-service performance monitoring can be realized through the achievement of the following goals:

- *Detection* of anomalous (potentially hazardous) changes in the transformer's internal condition
- *Diagnosis* of the present internal condition of the transformer based on detection of anomalies
- Determination of a *Prognosis* for the future behavior of the transformer based on past and present diagnoses

The goals of accurate in-service monitoring cannot, however, be met by the occasional observation of any single quantity. Rather, accurate and reliable monitoring can only be achieved through repeated *sensing of multiple quantities* in conjunction with the recognition of short-and long-term drifts, or *trends* in the *condition* of the transformer and its *signatures*. Additionally, the uniqueness of every transformer, even amongst a group of the same basic design, necessitates a monitoring scheme which is sufficiently intelligent to learn and interpret the characteristics of a particular transformer, that is, a scheme which *adapts*.

The problem of detection and diagnosis is further compounded by a general lack of knowledge concerning what really occurs in a transformer prior to failure; even if monitoring is possible there are many unknowns: what should be monitored and how often, what should be done with the accumulated data, how should the results be interpreted (what is normal, what is hazardous and may lead to failure), and what operator responses are appropriate given that a valid diagnosis is made?

The recognition of short and long term trends in the condition of a transformer first requires an understanding of what the normal conditions of a transformer and its signatures are. This understanding can only be achieved via monitoring experience with operating transformers; trends may be analyzed only after the normal condition of a transformer has been identified through the determination of *parameters* which characterize the *signatures* of the transformer and remain constant throughout the transformer's normal operating range. Short term trends will generally provide clear indications of changes which should raise flags to the system operator. Long term trends may be caused by acceptable aging or more slowly developing incipient failures. In both the short and long term cases, *trend analysis* provides for recognition of patterns of operation which deviate from the norm.

Once the normal conditions of a transformer and its signatures are understood, a machine can perform trend analysis to detect anomalies. The machine may even, in some cases, be able to diagnose the condition of the transformer; however, human input is probably necessary to develop a complete diagnosis and prognosis for the transformer's future.

This paper begins with a short description of the economic value of a transformer performance monitoring system. It then describes the structure of the Adaptive Transformer Monitoring System under development at MIT. This

[1]This work is led by James R. Melcher and Chathan M.Cooke and sponsored through the MIT Energy Laboratory Electric Utility Program by: Allegheny Power System, American Electric Power Service Corporation, Bonneville Power Administration, Boston Edison Company, Empire State Electric Energy Research Corporation, New York Power Authority, Northeast Utilities Service Company, Southern California Edison Company, Tokyo Electric Power Company, Inc. This support is gratefully acknowledged.

[2]Many members of the Laboratory for Electromagnetic and Electronic Systems have participated in the overall research program. As in any project involving a large team, it is difficult to isolate the parts of the whole; however, the authors wish to thank the following people who contributed directly to the work reported in this paper: J.R. Melcher, J.L. Kirtley Jr., D.P. Flagg, D.S. Archer, D. Singh, E. Frank, and M.C. Zaretsky.

monitoring system structure utilizes information both from observed (or learned) conditions in the transformer and human experts to identify potential failure modes. The paper next discusses proposed approaches to *automatic* detection and diagnosis of incipient failures, followed by a description of the implementation of an automatic detection system in hardware and software. (There is no discussion of an automatic diagnosis system as an expert system shell to perform automatic diagnosis has not yet been implemented.) Finally, results of ongoing tests carried out in the Pilot Transformer Test Facility at MIT are presented. These tests involve the characterization of several normal signatures and the detection of a simulated incipient failure through continuous on-line monitoring of an in-service transformer.

ECONOMIC VALUE OF MONITORING SYSTEMS

The upper bound of the amount that a utility should be willing to pay for a transformer monitoring system is its economic value, which can be determined by calculating the costs that a utility avoids by detecting and correcting a failure in the incipient stage; that is, before the failure becomes catastrophic. These avoided costs are the sum of two distinctly different components. The first component of value is the capital *replacement* cost of the transformer; given the assumption that a transformer lacking a monitoring system would be severely damaged by a failure and that the monitoring system detects an incipient failure in time for the utility to take the transformer off line, repair it and return it to service. The second component is based on system operating costs. Because transformers are expensive and have relatively low failure rates, utilities do not provide 100% backup. Where redundancy exists, it is system redundancy rather than hardware redundancy, e.g., the system as a whole is re-dispatched to reduce load flows through particular points during the period in which a transformer is repaired or changed out. In calculating the economic value of each of these components it is necessary to quantify the probability of failure, i.e. transformers failure rates are approximately 2% per year, and to consider standard economic/financial discounting rules on the time value of the investment in the monitoring system.

Transformer Replacement (Capital) Value

The economic value of the first component is relatively easily calculated as the replacement cost of the transformer minus any actual cost to repair the transformer. This component can vary between zero, in the case in which the monitoring system detects an incipient failure but that failure is not repairable, to the full value of the transformer itself. In the best case the incipient failure is minor but the potential consequences are catastrophic, such as a loose lead connection or loose winding wedges. An example of the latter case can be constructed using the following assumptions:

- The replacement cost of a transformer is $1,000,000.
- If a detectable incipient failure is allowed to progress, the transformer will be destroyed.
- The cost of repairing the transformer when the failure is detected in an incipient stage is extremely inexpensive relative to the replacement cost of the transformer (i.e., thousands of dollars, not hundreds of thousands).
- The transformer failure rate is 2% per year.
- The monitoring system is imperfect, and some failures are instantaneous, so only half of the actual failures will be detected.
- The expected life of a transformer is 40 years.
- The discount rate is 14%.

Given these assumptions, the maximum *annual* amount the utility should be willing to pay to avoid catastrophic failure of a transformer is $10,000. Given an expected life of 40 years and a discount rate of 14%, the present value of this annual investment over the life of the transformer is $81,000. Therefore, the value to the utility of detecting an incipient failure is $.08 per dollar of replacement cost. This represents the *highest capital value* that can be placed on a monitoring system. The lower bound is clearly zero since in the worst case the detection of an incipient failure only allows the transformer to be brought off line efficiently and then junked.

System Operating Value

The individual components of an electric power system are chosen and structured such that the system structure operates at maximum reliability and minimum cost. When a critical component fails the system keeps running (generally) but the cost structure changes. This is most easily seen on the generating side. When a transformer failure forces a low-operating-cost generator to come off line (e.g., a nuclear plant has a forced outage), other generators higher in the loading order pick up the slack, but at a *higher system operating cost.* The same argument can be made for the transmission system. Its components are designed to maintain system operations at a least cost level. When one component trips out, the system is re-dispatched to reduce load at or through a specific node in the system, again leading to a stable system, but at *higher system operating cost.*

The system operating value of a transformer is, therefore, a function of the location of the transformer in the system and the length of time the transformer is down. The value is measured in terms of the additional *system* costs that are incurred to avoid the bottleneck caused by the loss of the transformer. If a transformer happens to be a Generator Step-Up unit (GSU), the generator is unavailable until a spare is connected, or the transformer is replaced. This frequently takes a month. If the transformer is at a major substation, the load carried by the substation must be reduced for the length of time the transformer is out of service, unless there is redundancy.

The system value of a transformer monitoring device is estimated using the same logic as applied to calculating the capital value. In this case the capital value of the transformer is irrelevant. What is relevant is the increased cost in alternate system operation brought about by the need to re-dispatch the system. Again, the use of the extreme case provides an upper bound to the system value of a transformer monitoring system. The assumptions for the extreme case are:

- The transformer failure rate is 2% per year.
- The monitoring system is imperfect, and some failures are instantaneous, so only half of the actual failures will be detected.
- The transformer is a GSU for a base load generator.
- There is no spare transformer available.
- Replacement of the transformer requires 30 days.
- The expected life of a transformer is 40 years.
- The discount rate is 14%.

The EPRI-developed Regional Electric Utility for the Southeast Region of the United States [14] is used to perform the system cost valuation. This scale model system has installed capacity of 18,300 MW and a peak load of 15,000 MW with 5200 MW of nuclear base load and 9100 MW of coal. Monitoring systems are placed on the five GSU's at the nuclear plants and it is assumed that transformer outages per year are reduced to 1% as discussed above. The expected annual system savings per monitor on the five plants would be $140,000. The present value of this annual system savings over the expected 40 year life of the transformers would be $1.13 million *per transformer.*

This average system value amount reduces as a function of the number of monitoring systems that are applied to GSU's because the incremental value of the energy saved is reduced as monitoring systems are added to generators higher and higher in the loading order. At the upper end of the loading order, the peaking plants, the value is effectively zero.

Economic Value, Total

The total economic value is the *sum of the capital value and the system value.* What is clear is that for many large transformers the system value swamps the replacement value in absolute magnitude. For a $10 million GSU saved from a catastrophic failure and requiring only a short (hours) down time for repair of the detected incipient failure, the economic value of the monitoring system would be over $2 million.

The economic value of a transformer monitoring system is further enhanced if the installation of a monitoring system allows a utility to reduce the level of redundancy necessary to maintain satisfactory system reliability. For instance, many large generating plants use three single-phase transformers in the generator step-up application. To maintain reliability, many utilities install four transformers where only three are used, so that when one fails, a replacement can be quickly connected in its place. If a transformer monitoring system enhances the availability of the plant enough, the fourth transformer can be eliminated, reducing the capital cost of the generator step-up transformer(s) by 25%.

At other locations within the system, the value is reduced as a function of the costs that can be avoided by prevention of catastrophic failure. Site selection is important but it is clear that the potential value of transformer monitoring systems is extremely high when both the replacement (capital) and the system costs are considered.

CONCEPT AND STRUCTURE

Two issues must be addressed before an on-line transformer monitoring system can be designed and implemented. These are:

- Which *quantities* should be measured?
- How should a *failure* be defined and detected?

The determination of the quantities to be measured started with a detailed literature review and discussions with utility representatives and transformer manufacturers. The results of these actions led to the development of a set of structural hypotheses concerning the subsystems of a transformer and the manner in which specific measurable quantities might map into failure modes in each of the subsystems. The subsystems include the Tank, Bushings, Core, Windings, Insulation, Oil, Auxilaries, Tap Changers and Sensors. Figure 1 shows both the general decomposition and a specific example of the manner in which the effects of a through fault might be seen in some of these subsystems.

Development of the structure of Figure 1 led to the establishment of the goal of developing an integrated monitoring system as differentiated from developing only a set of independent, new and/or improved sensors.

Expansion of the concepts shown in Figure 1 into the concept of an integrated monitoring system allows the relation of typical transformer failure modes to observable quantities. A matrix of these relationships is given in Figure 2.

Once the development of an integrated transformer monitoring system was defined as a goal, the problem of detecting and diagnosing failures could be addressed.

Many monitoring schemes and systems employ the concept of setting thresholds for the normal limits of operation. Excursions from normal operation, and consequently potential failures, are detected when the threshold limits are exceeded. For example, a transformer may have several levels of threshold detection on its winding hot-spot temperature sensor. As each threshold is exceeded a corresponding message is sent to the operator and control system, whether that message be an alarm or a trip. With this scheme there is no information generated regarding how the transformer operated before the threshold(s) were exceeded. This is an inherent limit on sensitivity.

Sensitivity may be increased if the operation of the transformer is monitored and compared to normal *at all times*. This monitoring scheme, however, requires a better knowledge of *what is normal*. One way of achieving better knowledge of *normal* is to develop mathematical models for the normal operation of subsystems of the transformer, and compare the actual operation of those subsystems to the models in real time. This concept is presented in Figure 3. In Figure 3 any deviation from normal results in a non-zero error signal. The structure of the mathematical model of Figure 3 is chosen so that the *parameters* (or coefficients) of the model remain *constant* when the transformer is operating normally. The *parameters* then characterize a particular subsystem, or *signature* of the transformer.

The Module

The necessity of being able to *adapt* to a particular transformer is handled by estimating the parameters of the model using actual data from the transformer being monitored. Assuming that a given transformer is normal when new, (having passed its initial acceptance tests), the parameters of a model may be estimated *on-line*. The error term, called a *residual* then reflects the deviation of the transformer from its own normal state in the short-term, on the order of minutes-to-hours. If the parameters of a model are periodically re-estimated, on a daily or weekly basis, a long-term tracking (days-to-weeks) of the *condition* of that particular signature may be accomplished. These concepts of adaptability and short- and long-term tracking are embodied in the block diagram of a *module* given in Figure 4.

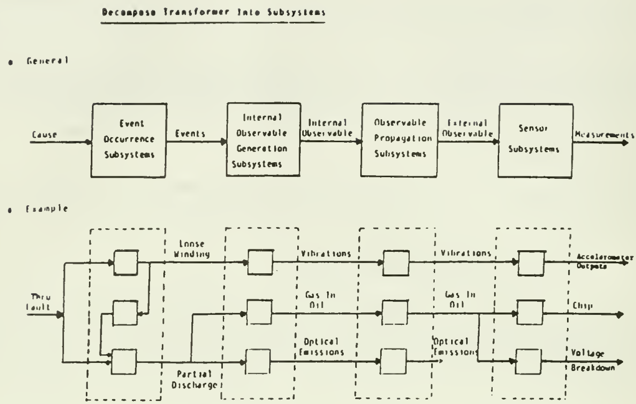


Figure 1: Decomposition of Transformer Failures

Failure Modes

Observable Quantities

	Bent Winding	Core Damage	Cracked Bushing	Electri-fication	Hot Spot	Arcing Short	Gas Bubbles	Contaminated Oil
Moisture-In-Oil			●		●			●
Gas-In-Oil			●	●	●	●	●	●
Partial Discharges			●	●	●	●	●	
Thermal	●	●		●	●			
Vibrations	●	●						
Oil Breakdown							●	●
Electrical					●			

Figure 2: Relationships Between Failure Modes and Observable Quantities

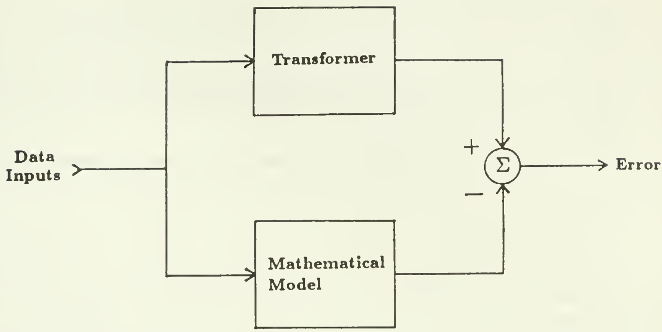


Figure 3: Monitoring System Fundamental Operation

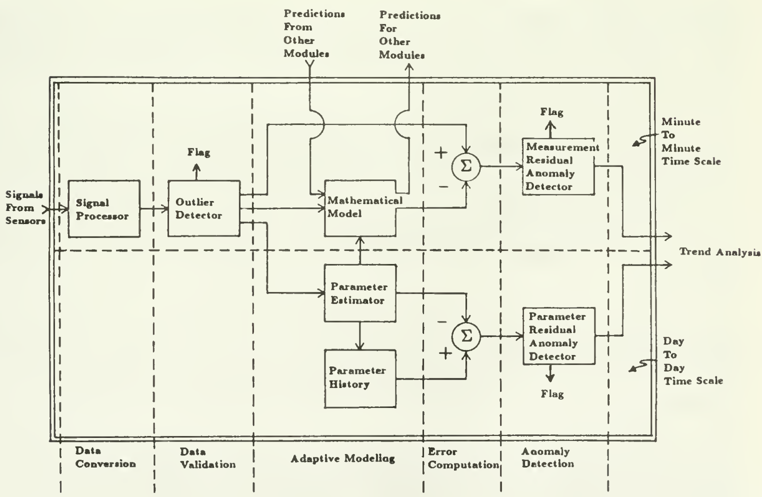


Figure 4: Module Block Diagram

A module [2] is implemented primarily in software. A list of definitions pertaining to Figure 4 is now given.

- Signals (data) from sensors pass to the *Signal Processor* where any necessary data preparation or reduction steps are performed.
- Processed data then moves to the *Outlier Detector* where threshold checks for bad data are made; bad data is announced to the human operator and the detection/diagnosis system with a Flag.
- Validated data is used as the input to a *Model* which predicts the values (of the Signature in question) that are expected during normal operation of the device being monitored. Additionally, the model may accept predictions from other modules as inputs and may output predictions for other modules. These additional inputs and outputs are used for compensation purposes, e.g., temperature compensation.
- Predicted values are compared to measured values in the *Measurement Residual Anomaly Detector*. This block looks for levels, rates-of-change, and patterns which are abnormal. If an abnormality is detected, the human operator and the detection/diagnosis system are alerted with a Flag.
- Periodically, the parameters (coefficients), of the mathematical equation which makes up the Model are updated, using measured values, through operation of the *Parameter Estimator* to assure that the Model remains accurate. When the Parameter Estimator operates, it automatically checks the new parameters for validity before installing them. (If the parameters are estimated using information-poor data, they will not accurately characterize the Signature). Valid parameters are also passed to the *Parameter History* for use in anomaly detection.
- The parameters of the Model are then tracked by the *Parameter Residual Anomaly Detector* to discriminate between acceptable changes, such as normal aging, and anomalies caused by incipient failures. As with the Measurement Residual Anomaly Detector, this block checks for anomalous levels, rates-of-change, and patterns. When an anomaly is detected, the human operator and the detection/diagnosis system are alerted.

The vertical dotted lines in Figure 4 divide the module up into five functional sections: Data Conversion, Data Validation, Adaptive Modeling, Error Computation, and Anomaly Detection. The horizontal dotted line divides the module according to time scales: the top half of the module operates on the Minutes-to-Hours time scale, and the bottom half operates on the Days-to-Weeks time scale.

In the intervals between installations of updated parameters (newly estimated parameters satisfy the parameter validity criteria), the condition of the signature and the accuracy of the model are checked via the measurement residuals. If the measurement residuals are small, the previously estimated parameters still accurately characterize the signature, and the condition of the signature is normal. If the measurement residuals exceed established limits (in level, rate-of-change, or pattern), an anomaly is detected even if the measurement residuals return to normal when a new set of valid parameters are installed. In this case, there has been a change in the condition of the signature, but the structure of the model still correctly describes the signature. If the measurement residuals exceed established limits and newly estimated parameters are systematically failing the validity test, the condition of the signature has changed so much that the structure of the model is itself no longer valid. This is another (probably more serious), form of anomaly.

Looking back at Figure 2, a one-to-one mapping can be made between *observable quantities*, *signatures*, and *modules*. A subset of the observable quantities listed in Figure 2 can be chosen as modules to provide the capability of detecting a majority of the failure modes listed.

The Monitoring System

A module exhibits increased sensitivity to incipient failures which affect the condition of a particular signature. This is due to the adaptive model and continuous real-time operation. Sensitivity to incipient failures can be increased even further by cross-correlating the detection outputs of various modules. To do this, it is necessary to combine these modules in a system which can *control and schedule* Data Acquisition, Information Organization, Module Operation, Detection, Diagnosis, Prognosis, Communications and Interfacing with the Operator.

The block diagram for such a system is given in Figure 5.

The system implemented in a combination of hardware and software, performing the functions listed above while mediating scheduling and data conflicts. The activities of system blocks include:

- Acquisition of raw data from *Sensors*
- Organization of raw data into a time-correlated format in the *Primary Buffer*, thus making the raw data available to the remainder of the system
- Processing raw data in *Modules* to extract information relevant to a determination of whether or not the transformer being monitored is operating normally
- Placement of relevant information from modules into the *Secondary Buffer*, for use by the rest of the system
- Performance of *Trend Analysis* on raw data and relevant information from modules to *Detect* anomalies in the transformer being monitored, *Diagnose* the condition of the transformer, and deliver a *Prognosis* on the future operation of the transformer
- *Organize and Schedule* all of the above, and provide operator interface, through the operation of a *Controller*

In summary, the MIT-developed monitoring structure is an integrated system with the Module as its core. Conceptually, each of the functions of the system operate independently and *in parallel*, sharing information when required. This functionality permits the overall system to be highly flexible. Since information organization and scheduling of operations are handled by the system, resulting in a well-defined interface between modules and the system, modules may be added or removed easily. The final block in Figure 5, *Trend Analysis*, integrates the information flows from the individual modules to provide the knowledge upon which diagnostics can be based.

Trend Analysis

Trend Analysis is the final step in the process of transformer monitoring. The MIT project has defined the structure of trend analysis, but to date, has not fully implemented that structure. The discussion which follows provides the specifications for implementation of trend analysis given available module data.

As outlined in Section , accurate in-service performance monitoring of transformers can be realized with the achievement of three goals:

- *Detection* of anomalous (potentially hazardous) changes in the transformer's internal condition
- *Diagnosis* of the present internal condition of the transformer based on detection of anomalies
- Determination of a *Prognosis* for the future behavior of the transformer based on past and present diagnoses

Trend Analysis is involved with achieving all three of these goals. The first two goals are near-term in the sequence of system development, in fact, they are very much intertwined; the third is somewhat farther down the road as it requires substantial experience with on-line monitoring to achieve.

Detection. *Detection of anomalous change* is split between individual Modules and the Trend Analysis block. As described above, a Module tracks trends in an individual signature, automatically and independently detecting anomalies in that signature. The Trend Analysis block automatically detects anomalous changes in the transformer by *cross-correlating* trends and anomalies *between modules.* As with module-level testing, system-level testing concentrates on levels, rates-of-change, and patterns which are abnormal. This cross-correlation carries over into diagnosis, as discussed below.

In this approach to transformer monitoring, sensors are considered a subsystem of the transformer. As such, failure of a sensor is treated as a failure of the transformer, albeit a generally non-critical failure from the operator's point of view. From the system's point of view, failure of a sensor will cause the module using that sensor to detect an anomaly, in the same manner as detection of a failure in one of the transformer's other subsystems. Sensor failure/bad data is detected using standard procedures which have been successful in other applications [3,4,5]. Three types of bad data are hypothesized:

- Intermittent Failure: Usually good but bad sometimes
- Jump Failure: Suddenly bad all the time
- Drift and Offset Failure: A steady or increasing bias

Hypothesis testing techniques are used to determine if the above-listed hypotheses can account for the detection flags raised by modules. Diagnosis of bad data to determine bad sensors is possible only if there is enough *cross sensor redundancy* built into the monitoring system. Even though this redundancy does not necessarily require multiple units of the same type of sensor, but rather, the knowledge to determine if a particular sensor has failed using the information contained in signals from various sensor types, it is expected that diagnosis of some types of sensor failures will not be possible without human help.

Sensor failure is detected in this section of the system because the overhead involved with the operation of this bad data/sensor detection system on the front end of the monitoring system would make continuous on-line monitoring much more difficult to achieve; a possible future goal is to utilize sensors which are smart enough to detect and diagnose self-failures in real time, thereby relieving the monitoring system of this burden.

Diagnosis. The diagnosis function of trend analysis tries to determine the reason(s) for any anomalous behavior that is detected. Diagnosis is more difficult than detection. The initial phase of the diagnosis operation will be performed automatically. (A human expert may simply accept the result of the automatic operation, or use it in an effort to arrive at a more complete diagnosis.) Anomalies, as discussed in Section are the primary stimulus for automatic diagnosis. They are not, however, the exclusive inputs to the diagnosis operation. The Trend Analysis block *diagnoses the transformer's condition* (including full or partial diagnosis of bad sensors), based on all the information available to the system: detected anomalies, trends in measurements and parameters, and trends in measurement and parameter residuals. For instance, trends that have not been flagged as anomalous may influence a particular diagnosis. It is for this reason that the cross-correlation of information from multiple signatures is important, e.g., a slight trend in a parameter associated with one signature may be significant in the presence of anomalous behavior in a second signature.

Tests performed in the diagnosis stage involve the relation of current information to particular failure modes. It is, however, conceivable that having detected an abnormal condition in the transformer, the system may not possess enough evidence to reach a conclusive diagnosis. In this case, the cost of mis-diagnosing possible failures must be weighed against the consequences of continued operation of the transformer. A remedial action in this situation may be the initiation of more costly tests. One such test is the performance of a dissolved gas analysis on a manually-drawn oil sample, the results of which are used as further input to the diagnosis system. (Before requesting this action, the expert system will weigh the cost of sending out the technician and the probable amount of information to be gained by the test, against the uncertainty in the diagnosis.) With this new information, the expert system may be able to arrive at a diagnosis.

The relation of current information to particular failure modes likely involves linear or nonlinear combinations of the information associated with several signatures. Some of these combinations can be explicitly specified using knowledge available today; e.g., there is a large body of information available concerning dissolved gas analysis. However, for many of the signatures monitored by the prototype MIT system (these signatures are described below), it is not yet possible to specify explicit tests, particularly for combinations of signatures. This uncertainty is based on the fact that the necessary data is not yet available. The knowledge base required for the diagnosis system is being broadened with the MIT Pilot Transformer Test Facility and from field studies as prototype and field demonstration systems are installed and operated on other transformers[3]. Not enough is known about residual and parameter behavior in the face of specific incipient failures to project at what point particular diagnoses can be reached during the evolution of a failure.

Preliminary results, reported in Section , generate confidence that incipient failures can be detected before serious damage has occurred. With human interaction, the system will diagnose incipient failures long before traditional threshold techniques have enabled detection.

[3] This work is being commercialized by J.W. Harley, Inc., of Twinsburg, Ohio, and Westinghouse Electric Corporation's Materials and Manufacturing Technology Laboratory in Sharon, Pennsylvania.

It must be remembered that, with regard to diagnosis, this system is meant to be a tool which augments the abilities of the human expert.

Prognosis. Finally, the Trend Analysis block involves the development of a *prognosis for the transformer's future health.* That is, to decide whether or not the condition of the transformer is unsatisfactory, and what the probability of more severe failure is under various forms of continued operation (e.g., full or partial reduced loading). The prognosis function can be aided by the use of an expert system but the final decision will usually be based on human judgement.

In summary, the process of trend analysis is one based on the modular structure of the monitoring system. It builds on the output of the individual modules to identify changes in combinations of parameters and measurements that point toward incipient failure; and, in the final analysis, the potential cause of that failure. Trend analysis as a process will complement human knowledge-not replace it, in evaluating the condition of the transformer. It provides a continuous observation function, and an information resource not previously available to the decision maker.

IMPLEMENTATION

This section describes the implementation of a Pilot Monitoring System using the structure and concepts discussed in Sections , , , and . The Pilot Monitoring System developed by MIT is installed in the Pilot Transformer Test Facility in MIT's Building N10. It is a combination of computer hardware and software designed to fulfill the dual functions of: data acquisition for model and module development and implementation of an on-line transformer monitoring system. The discussion will first introduce the Pilot Transformer Test Facility, then present a more detailed system block diagram, and finally will proceed into a description of the actual hardware and software.

Pilot Transformer Test Facility

The center of the pilot facility is a 50 kVA, 240/8000 Volt, Single Phase, oil-filled, pole-type transformer. This transformer is known as the *Test Transformer.* The tank and transformer have been modified with the installation of numerous sensors; the tank does, however, retain its original gas space (sealed to the atmosphere and filled with dry nitrogen). The transformer has also been provided with a forced-oil circulation system to allow external control of heating and cooling. Excitation voltage and load current can be set independently. The Test Transformer is connected in parallel with a second, identical pole-type transformer. Variable loading to 150% of rated current at full voltage is achieved by using a third, smaller transformer to inductively drive circulating current through the two pole-type transformers. By controlling the phase of the circulating current, the Test Transformer may be made to look as if it is supplying real and reactive power to a load.

The 50 kVA size units were chosen to be large enough to have space for the needed sensors and to generate substantial core and winding losses during load cycles; yet small enough to allow easily-made changes to the monitoring structure, as well as to fit inside the laboratory building.

Pilot Monitoring System Structure

An implementation of the monitoring system discussed in Section involves more detail than presented in the structural diagram of Figure 5. This added detail, involving data and control paths, peripherals, and external communications, is depicted in the block diagram of Figure 6. The blocks in this system diagram are chosen to represent functional pieces of the Pilot Monitoring System; as such, some of the blocks represent hardware, some represent software, and some represent combinations of hardware and software.

The original goal was to implement a monitoring system on a personal computer. It became clear, however, as the Pilot Monitoring System was designed, that some sort of multi-tasking, multi-processing computer environment was necessary. The tasks to be executed, from data acquisition on microsecond-time-scales to parameter estimation on a daily-time-scale required more computational power and flexibility than one personal computer was capable of delivering. Consequently, a basic hardware structure of two IBM AT-compatible personal computers was settled on.

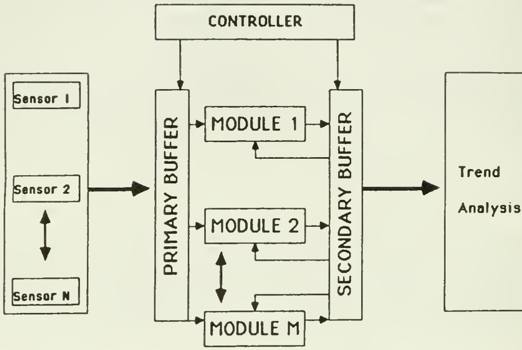


Figure 5: System Block Diagram

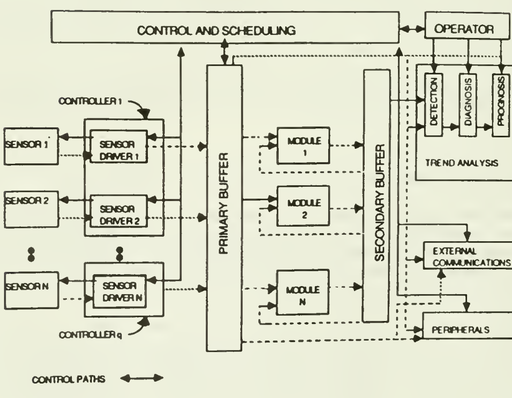


Figure 6: Pilot Monitoring System Block Diagram

Hardware Overview

The heart of the Pilot Monitoring System is a single IBM AT compatible machine running at 8 mHz under the IBM Xenix (Version 2.0) operating system. (Xenix is a version of UNIX.) This machine provides a multi-user, multi-tasking environment for the co-ordination and control of a data acquisition subsystem as well as processing the resulting data. This *Master Machine* has a number of peripherals attached to it including a printer, modem, color monitor, dual 20 megabyte fixed disk drives, 9 track open reel tape drive, dual floppy drives, 1 additional user terminal (with provisions for other serial devices), as well as a data acquisition subsystem.

The data acquisition subsystem is another IBM AT compatible machine, running at 6 mHz under MS-DOS 3.10 and coupled to a Keithley Data Acquisition and Control - Series 500 Measurement and Control System. The AT compatible, called the *Acquisition Machine* has a 20 megabyte fixed disk drive, dual floppy drives, an EGA video card, and monochrome video display. The system board has its memory split into two 512k blocks. The first block is used as DOS base memory. The second block is addressed above the system ROMs as extended memory and is used for a RAM disk. Other than drive controller and video display adapter, the only additional board in the expansion bus is the interface to the Keithley System 500 modular data acquisition system. This combination is responsible for obtaining temperatures from 23 thermocouples, vibration signals from 2 accelerometers, high and low side current and voltage wave forms and RMS values, and dissolved gas ppm from a Syprotec H-201R Hydran monitor. This subsystem is controlled by the master machine using an RS-232 serial line. Data is transmitted in batch every few minutes from the Acquisition Machine to the Master Machine over a second RS-232 line.

All of the analog data acquisition portion of the Pilot Transformer Monitoring System (data being acquired from the Pilot Facility Test Transformer) is handled by the above-mentioned Keithley Series 500 System operating in conjunction with the Acquisition Machine. The Keithley System consists of a self-contained chassis and motherboard with slots to accommodate ten (10) plug-in circuit boards. The slots accept a variety of boards designed to perform various data input and output, or control functions. The data acquisition chassis interfaces with the Acquisition Machine through a cable (or an MIT developed optic link) which connects to the interface card plugged into one of the Acquisition Machine's expansion slots.

This particular data acquisition system was chosen because of its extreme versatility, large number of available channels, and superior temperature measurement circuitry.

The combination of the Master Machine, Acquisition Machine, and Keithley System forms a loosely-coupled multi-tasking, multi-processing computer system.

Acquisition Machine Software

Operation of the Keithley System 500 is through software running on the Acquisition Machine. This software is a combination of commercial and custom written code. Fundamental operation of the System 500 is performed by a software package supplied by Keithley. This package is called SOFT500, and it operates as a superset of commands in the interpretive BASIC language environment. The data acquisition routines, or drivers, are therefore, custom-written BASIC programs with imbedded SOFT500 commands.

Data acquired by the System 500/Acquisition Machine combination is pre-processed in the Acquisition Machine to cut down on the data transfer requirements of the overall monitoring system. Pre-processing involves computation of RMS values, averaging, scaling, and other data reduction operations. Pre-processing is done with compiled routines written in C to increase computation speed and aid portability. After pre-processing, the reduced data is transferred to the Master Machine for further processing and analysis.

Master Machine Operating System

The operating system chosen for the Master Machine is UNIX. UNIX is a well-established multi-tasking operating system developed by A.T. & T. Bell Labs. The current version is UNIX System V. It is available on many different computers and provides good support for the C programming language. The wide availability of UNIX System V and C means that software written in C or imbedded with UNIX system commands is not restricted to one computer. If written properly, the software is quite portable. Furthermore, UNIX contains many system commands useful to the Pilot Monitoring System, and is based on a file system structure which easily lends itself to the buffering and shared information demanded by the monitoring system.

The version of UNIX chosen for the Pilot Monitoring System is IBM Xenix (Version 2.0). IBM Xenix was picked because, among several UNIX operating systems available for AT's and compatibles at the time of selection (1987), it was the only system with proven reliability.

Master Machine Software

The specifications for the monitoring system call for a coordinating element to synchronize the activities of the individual modules. The operation of this coordinating element is required to be independent of the particular actions a module performs and, in fact, independent of the number of modules being coordinated. The specification also calls for the establishment of a mechanism for passing data between various modules, while limiting the constraints on the number and types of modules running. This mechanism will perform the duties of the primary and secondary buffers in the system block diagram.

Together these two requirements necessitate a standardized interface for the modules. It was decided that a module would only be required to perform a given set of actions at a pre-defined interval. The module would then respond to some trigger from the coordinating element by performing this set of actions, secure in the assumption that the module is synchronized with the system.

For flexibility, each module may also have its own initialization and/or termination code. The initialization code is triggered simply by starting the module. If the initialization fails, the normal trigger is taken as an initialization trigger until it succeeds. There is a separate termination trigger that causes termination code to be executed. The termination code will be executed after the normal set of module actions until it succeeds, at which time the module exits.

Inter-module communication of data is handled through the file system of the host computer. A limited buffer is provided for efficient retrieval of recent data.

Dispatch Software

The coordinating element consists of a single process that coordinates an arbitrary number of individually compiled programs. The resulting process is alternately referred to as *dispatch*, the scheduler or the synchronization process.

The programs which are coordinated by the synchronization process are referred to as modules. These modules are implemented specifically to fit into this scheme. (The structure of a module is discussed in Section . Each module is a separately compiled program. Because of this, the set of presently executing modules can be modified with ease and the addition of new modules has little or no impact on existing modules. The set of modules which is to be run is established through the use of an input file, also referred to as the jobs file. The modules run continuously in the background and are triggered to execute various portions of their code by the synchronization process. Dispatch can determine the execution status of each module and, if a module is not ready to be triggered at the appropriate time, a count of missed intervals would be incremented. When the module is ready to be triggered, it may perform some processing based on this value. In this way, each module is kept synchronized with the entire system.

Module Software

From a software point of view, a module consists of four parts: an initialization routine, a normal iteration routine, a synchronization error recovery routine and a termination routine. Though a module is a separately-executable program, it must be run by a synchronization program to operate correctly. A set of module utilities have been provided to interface the module with the dispatch process.

MIT chose to develop modules for the following signatures:

- Thermal (IEEE Loading Guide Model)
- Thermal (Constrained Flow Model)

- Winding Vibration (Black-Box Model)
- Dissolved Gas In Oil (Thermal Based Model)
- Dissolved Moisture In Oil (Thermal Based Model)
- Partial Discharges (Electrically Based Model)

Unfortunately, not enough progress was made on the development of an electrically-based sensing-scheme for partial discharge detection to warrant development of a module; therefore, partial discharges will not receive further consideration in this paper.

The present status of the remaining five modules will now be discussed. In the interests of space, detailed discussions of the models contained in each module will be omitted. References will be listed, however [6].

Thermal Module (IEEE Loading Guide Model): Thie3mod. One purpose of this module is to detect changes in the thermal system of the transformer, particularly excess heating. A second purpose is to predict un-measurable temperatures to be used in compensating the models in other modules (e.g., dissolved gas module). A third, as-yet-unrealized purpose is to enhance loadability by running the model faster than real time to allow the operator to foresee the consequences of operational decisions (e.g., overloading during peak periods).

This module is based on the IEEE/ANSI Loading Guide Models for prediction of top oil temperature and hot spot temperature using ambient temperature and load current as inputs [7]. The standard models have been modified to allow the top oil model to adapt to the transformer on-line [8], parameter estimation is performed using the Least Squares Method; the hot spot model is not adaptive, relying on parameters measured during initial heat runs:

- Measured ambient temperature and load current are used to predict top oil temperature; dynamic model (every two minutes)
- Measured top oil temperature is compared to the top oil temperature prediction to calculate a measurement residual with level detection (every two minutes)
- Measured top oil temperature and load current are used to predict hot spot temperature; static model (every two minutes)
- Top oil temperature predictor parameters are estimated using load current and measured ambient and top oil temperatures (every 24 hours)
- Top oil temperature predictor parameters are tracked graphically
- Winding internal temperature prediction is used as a compensating input to a winding vibration module

Thermal Module (Constrained Flow Model): Thmod. One purpose of this module is to detect changes in the thermal system of the transformer, particularly excess heating. A second purpose is to predict un-measurable temperatures to be used in compensating the models in other modules (e.g., winding vibration module). A third, as-yet-unrealized purpose is to enhance loadability by running the model faster than real time to allow the operator to foresee the consequences of operational decisions (e.g., overloading during peak periods).

This module uses more accurate models than the IEEE module; physically-based equations have been developed to predict temperatures in and near regions of constrained oil flow, such as cooling ducts in windings, and at locations in the winding bulk [8]. More dynamics are included than in the IEEE models. Three ducts have been instrumented in the Test Transformer: one specifically constructed for the purposes of experimentation called the artificial duct, and two actual ducts in the high voltage section of the winding, arbitrarily designated the thermocouple-side duct and the accelerometer-side duct. The disadvantage to this module is that it requires oil temperature measurements to be made in regions near the winding, although not actually inside the winding. The models which predict oil temperatures are adaptive, the models which predict winding surface and internal temperatures are partially adaptive. Parameters are estimated using the Least Squares method:

- Measured duct bottom (inlet) oil temperature and load current are used to predict duct top oil (outlet) temperature; dynamic model (every two minutes)
- Measured duct top oil temperature is compared to the duct top oil temperature prediction to calculate a measurement residual with level detection (every two minutes)

- Measured duct top and bottom oil temperatures and load current are used to predict oil temperature at any location within a duct; dynamic model (every two minutes)
- Predicted duct internal oil temperature and load current are used to predict winding surface temperature; static model (every two minutes)
- Predicted winding surface temperature and load current are used to predict winding internal temperature; dynamic model (every two minutes)
- Duct top oil temperature predictor parameters are estimated using load current and measured duct top oil temperatures (every 24 hours)
- Duct top oil temperature predictor parameters are tracked graphically
- Hot spot temperature prediction is used as an input to a thermally-based dissolved gas module

Winding Vibration Module (Black-Box Model): Vibmod. The purpose of this module is to detect potentially dangerous changes in the physical structure of the winding (e.g., loose wedges) caused by events such as through faults.

This module uses as its inputs: a core vibration time series signal acquired from an accelerometer mounted on the core, a winding current time series signal taken from a current transformer (CT) on the low voltage side which is squared in software, RMS terminal voltage, predicted winding internal temperature. The module performs a Fourier transform on the time series core vibration and load current squared data. The complex Fourier coefficients for the first three harmonics of these signals are input to a black-box model. Based on these inputs the model predicts the Fourier coefficients of the first three harmonics of the winding vibration. The model contains no dynamics but is completely adaptive. The predicted winding vibration Fourier coefficients are compared to measured winding vibration Fourier coefficients (calculated using a time series signal acquired from an accelerometer mounted on the winding) and a measurement residual is computed. Parameters are estimated using the Least Squares Method [9,10,11,12]:

- Time series data is acquired from load current CT, core accelerometer, and winding accelerometer. Complex Fourier transforms of each signal are performed (every 10 minutes)
- Load current squared and core vibration harmonics, RMS terminal voltage, and predicted winding internal temperature are used to predict winding vibration harmonics; static model (every 10 minutes)
- Measured and predicted winding vibration harmonics are compared to compute a winding vibration measurement residual with level detection (every 10 minutes)
- Winding vibration predictor parameters are estimated using measured winding vibration, measured core vibration, load current squared, terminal voltage, and predicted winding internal temperature (when enough data to estimate good parameters becomes available)
- Parameters are tracked graphically

Dissolved Gas In Oil Module (Thermal Based Model): Gasmod. The purpose of this module is to detect anomalous changes in the dissolved gas content of the oil. The model is partially black-box, partially physically-based, and is intended for use with the Syprotec H-201R Hydran Dissolved Gas Monitor. The Hydran is sensitive to Hydrogen, Carbon Monoxide, Acetylene, and Ethylene. The module actually runs two models, both predicting the dissolved gas reading of the Hydran. One model uses measured top oil temperature as its input, the other model uses predicted hot spot temperature as its input. The models are static and adaptive. Parameters are estimated using the Least Squares Method:

- Measured top oil temperature and predicted hot spot temperature are used to make two separate predictions of the Hydran dissolved gas reading; static models (every 10 minutes)
- Predicted Hydran readings are compared with actual Hydran measurements to compute dissolved gas measurement residuals with level detection (every 10 minutes)
- Model parameters are estimated using measured top oil temperatures and Hydran readings for one model and predicted hot spot temperature and Hydran readings for the other (every 24 hours)

- Model parameters are tracked graphically

Dissolved Moisture In Oil Module (Thermal Based Model): Wthmod. The purpose of this module is to detect anomalous changes in the dissolved moisture content of the oil. Such changes (usually an increase) indicate deterioration of the paper insulation due excessive heating and/or acid attack.

This module computes an approximation of oil moisture content based on a temperature reading. Again, two models are running, one based on top oil temperature, and one based on hot spot temperature [13]. Presently, no residual is calculated on-line, due to the lack of availability of a solid state moisture sensor. Moisture readings are therefore made by hand, as is the measurement residual calculation. The models are static and adaptive. When on-line measurements become available, parameters will be automatically estimated using the Least Squares Method:

- Measured top oil temperature and predicted hot spot temperature are used to make two separate predictions of the dissolved moisture reading; static models (every 10 minutes)
- Predicted moisture readings are compared by hand with actual moisture measurements (Karl-Fischer Method) to compute dissolved moisture measurement residuals (every 5 days)
- Model parameters are estimated using measured top oil temperatures and moisture readings for one model and predicted hot spot temperature and moisture readings for the other (every 2 months)
- Model parameters are tracked graphically

Module and System Summary. The dispatch process and the module interface have proven to be a flexible mechanism for implementing the various modules. The dispatch process is independent of the functions of the modules under its control. As such, bringing a new or updated module on line is simply a matter of editing an input file to reflect the new set of modules (and their schedules) and re-invoking the dispatch process. Communication between the dispatch process and an individual module follows the same lines regardless of the particular module being driven, modified only by the schedule provided in the input file.

Using the module interface reduces the problem of implementing a new module to implementing just those routines that distinguish one module from another. In effect, one just implements the mathematical model at the heart of the module. All problems of scheduling and communication have been abstracted away.

Each individual module is designed to capture the function of some subsystem of the transformer. Thie3mod and Thmod handle the thermal system, Vibmod deals with the windings, and Gasmod and Wthmod handle the oil and insulation systems. In describing the function of a transformer subsystem, each module embodies a mathematical model of how that system works. The mathematical model may be intended to describe a physical model, such as the Thmod's constrained flow model, or may describe an observed functional relationship, such as in the Wthmod (moisture module). In either case, the mathematical model contains parameters that adapt to observed conditions, to *tune* the module to the actual behavior of the transformer. The design of the module system is intended to simplify the process of inserting a particular model into the system and allow for the maintenance of the adaptive parameters.

EXPERIMENTAL RESULTS

This section presents experimental results from the MIT Pilot Transformer Test Facility. Included are plots of normal module operation and plots of residual behavior during a simulated failure - unexpected dissipation of heat in the transformer's oil space. Note: Whenever labels at the top of plots contain arrows, the arrows indicate which vertical axis is associated with that particular data.

The first data presented characterizes normal load cycle operation of the Test Transformer. Figure 7 shows the low-side voltage and current for a period of three days. Rated voltage is 240 Volts and rated current is 208 Amps. The dip to zero in the voltage and current on 4/3/89 indicates the transformer was shut down briefly to draw an oil sample.

Figure 8 shows operation of the constrained flow thermal module over the same period of time. Note the residual, Curve A, oscillates about zero indicating good agreement between measured and predicted values.

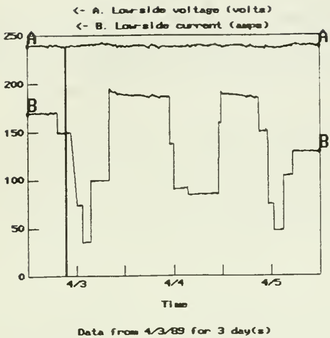


Figure 7: Normal Load Cycling

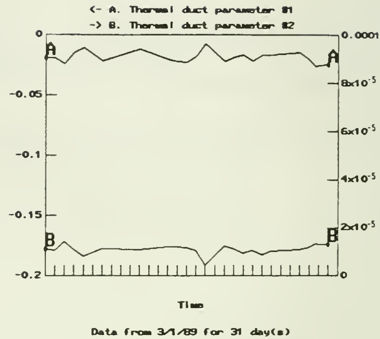


Figure 9: Constrained Flow Parameters

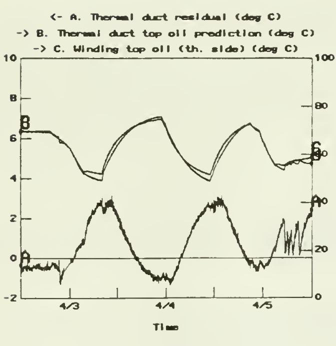


Figure 8: Constrained Flow Thermal Module

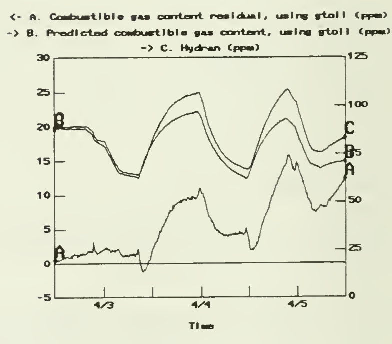


Figure 10: Dissolved Gas Module

Figure 9 is also from the constrained flow thermal module, indicating a month's worth of parameters. It is seen that the parameters are quite stable, no changes have occurred in the condition of the transformer.

Figure 10 depicts the dissolved gas module, again for April 3-5, 1989. The combustible gas content is oscillating with temperature around 20 ppm. The residual is on the order of 5-10 ppm.

Figure 11 shows a hand calculation of the moisture residual over a three month period. Due to the lack of a functioning on-line moisture sensor, this moisture monitoring is done completely by hand, using oil samples drawn every few days from the Test Transformer. However, even with infrequent sampling, it is seen that the moisture model (based on oil temperature) is quite accurate, and the moisture content of the transformer has not changed significantly during the period shown.

The next three plots depict operation when a simulated failure was introduced into the transformer in the form of unexpected heating. While the transformer was operating in steady-state at 75% of full load, as indicated by Figure 12, a heating tape was used to inject approximately 30 Watts of heat into the side of the transformer's tank. This amount of heating is equivalent to about 10% of the losses of the transformer.

It is seen in Figure 13 that the combustible gas residual undergoes a step change to a very high value. This is because the heating tape was disturbing the dissolved gas sensor.

Figure 14 shows a corresponding increase in the constrained flow thermal residual. The model predictions are no longer accurate because there is heat appearing in the tank which is not due to normal load losses.

In this example, the dissolved gas and thermal modules have both detected anomalies. In one case, the anomaly is due to a type of sensor failure (the temperature compensation of the gas sensor was impaired). In the other case, the anomaly is in the thermal signature of the transformer. This example serves to show that the monitoring scheme presented in this paper can *detect* anomalies. In fact, the distinct step in the thermal residual has a magnitude of approximately one degree. This means that the oil temperature in the transformer was *one degree above normal.* Standard threshold alarms would not have caught an incipient heating failure until the excess heating was much worse.

CONCLUSIONS

An economic argument for the installation of transformer performance monitoring systems on large power transformers has been given. A scheme for on-line performance monitoring of large power transformers has been presented. A relatively inexpensive prototype laboratory implementation of the monitoring scheme (lacking an expert system shell to perform diagnosis) has been described. Finally, results indicating the sensitivity of the monitoring scheme to an incipient failure have been presented, showing that the system is much more sensitive than standard threshold level detection.

Additionally, it should be noted that this monitoring system is not limited to the modules and sensors described in this paper. There is ongoing research at MIT, and elsewhere, directed toward the development of new sensors and modules. These new sensors and modules can and will be readily accomodated.

APPENDIX IEEE THERMAL MODULE

The description which follows summarizes the functions being performed by the IEEE thermal module. The equations used have been drawn from the IEEE loading guide[7] and manipulated into discrete-time form. This description is representative of the detail required for each module in the system.

Model The model being implemented is

$$\begin{aligned} pgtoil[k] &= A * (pgtoil[k-1] - gambient[k-1]) + \\ & B * ilow[k]^{1.6} + \\ & gambient[k], \end{aligned}$$

where *pgtoil* is the predicted mixed top oil temperature, *gambient* is the ambient temperature, and

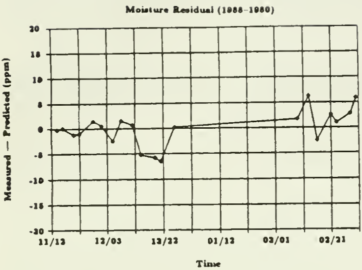


Figure 11: Dissolved Moisture Module

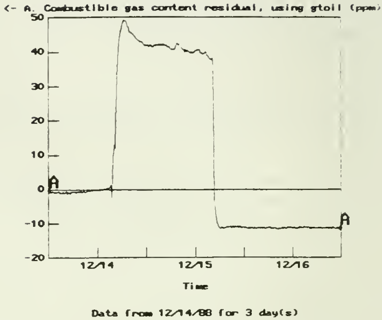


Figure 13: Anomalous Gas Residual

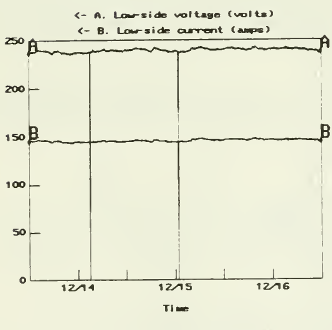


Figure 12: Steady-State Operation

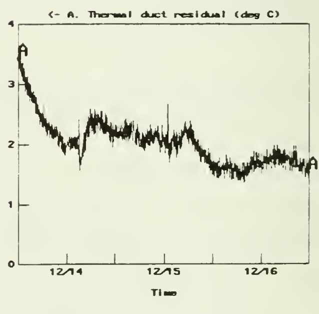


Figure 14: Anomalous Thermal Residual

ilow is the load current. A and B are adaptive parameters which are periodically re-estimated.

The IEEE thermal module generates a prediction of hot spot temperature for other modules to use for temperature compensation. The equation used is

$$pwtint[k] \quad = \quad C * ilow[k]^{1.6} + gtoil[k],$$

where C is an inestimable parameter calculated during the initial heat run, and *gtoil* is the measured mixed top oil temperature.

The initial prediction of mixed top oil temperature is set equal to the initial reading of mixed top oil temperature ($pgtoil[0] = gtoil[0]$). The model used to calculate *pwtint* is static, so no special initialization is required.

Outlier detector The inputs are checked against operator-specified limits. If these limits are violated, the operator is notified. Presently, these limits are simple thresholds specifying a valid range of inputs and/or a maximum rate of change from one instance to the next.

Measurement residual anomaly detector Measurement residual anomaly threshold detection is handled in a manner similar to outlier detection. A valid range of residual values and a maximum rate of change can be specified by the operator. The residual in this case is

$$rgtoil[k] = gtoil[k] - pgtoil[k],$$

where *rgtoil* is referred to as the mixed top oil temperature residual.

Parameter estimator The equation used to estimate the parameters for the module is

$$\begin{aligned} gtoil[k] \quad - \quad & gambient[k] = \\ & A * (gtoil[k-1] - gambient[k-1]) + \\ & B * ilow[k]^{1.6}, \end{aligned}$$

using a least-squares algorithm.

Note that the actual measured mixed top oil temperature (*gtoil*) is used to generate the parameters, thus adapting the model to the (possibly changing) internal condition of the transformer.

At present, parameters are re-estimated daily using two days worth of data. Operator experience is used to establish thresholds to screen out parameters estimated from information-poor data. This threshold is compared to a number generated by the estimation routine that remains small only when the new parameters yield a good curve fit and the input to the estimation routine is well-conditioned (information-rich).

Parameter residual anomaly detector Parameters, like input data and measurement residuals, are compared to operator-specified limits for value and rate of change. Again, the operator is notified of any anomalies.

REFERENCES

1. D. J. Allan, "Editorial," Electra, No. 114, pp. 3-6, October 1987.
2. F. C. Schweppe, "In Service Transformer Monitoring," Internal Report, M.I.T. Laboratory for Electromagnetic and Electronic Systems, December 1986.
3. F. C. Schweppe, Uncertain Dynamic Systems. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1973.
4. L. Ljung, System Identification: Theory for the User. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1987, ch. 14, pp. 386-389.
5. L. F. Pau, Failure Diagnosis and Performance Monitoring. New York: Marcel Dekker, Inc., 1981.
6. J. R. Melcher and C. M. Cooke, Co-Principal Investigators, Trend Analysis: Performance Monitoring of Transformers, Final Report, Laboratory for Electromagnetic and Electronic Systems and Energy Laboratory Electric Utility Program, Massachusetts Institute of Technology, Cambridge, Massachusetts, October 1988.
7. "Guide for Loading Oil Immersed Distribution and Power Transformers Rated in Excess of 100 MVA," IEEE PES Standards, July 1982.

8. D. S. Archer, "An Adaptive Thermal Module for Transformer Monitoring" S.M. Thesis, M.I.T. Department of Electrical Engineering and Computer Science, August 1988.

9. J. C. Lavalle, "Failure Detection in Transformers Using Vibrational Analysis," S.M. Thesis, M.I.T. Department of Electrical Engineering and Computer Science, September 1986.

10. D. J. McCarthy, "An Adaptive Model for Vibrational Monitoring of Power Transformers," S.M. Thesis, M.I.T. Department of Electrical Engineering and Computer Science, May 1987.

11. J. L. Kirtley Jr., J. C. Lavalle, and D. J. McCarthy, "Acoustic Monitoring of Transformer Structures." Paper No. 820-03 presented at CIGRE Symposium on Monitoring for HV Systems, Vienna, May 1987.

12. D. Singh, "The Impact of the Political and Economic Environment on Policy Choices - The Case of the Electric Power System in the United States and India," S.M. Thesis, M.I.T. Technology and Policy Program, February 1989.

13. M. C. Zaretsky, "Parameter Estimation Using Microdielectrometry With Application to Transformer Monitoring," Sc.D. Thesis, M.I.T. Department of Electrical Engineering and Computer Science, November 1987.

14. EPRI, The EPRI Regional Systems, EPRI Special Report #P-19550-SR, Palo Alto, CA, July 1981, 148 pages.

TOGA™ (Transformer Oil Gas Analyst): The Evolution of an Expert System

JILL RAYE HOWES
The Hartford Steam Boiler Inspection and Insurance Company
One State Street
Hartford, Connecticut 06102, USA

ABSTRACT

TOGA, the Transformer Oil Gas Analyst, is an expert system that identifies incipient faults in oil-cooled transformers and analyzes the condition of the insulating oil. It examines data from both oil and screen tests and recommends when the transformer should be resampled.

TOGA is part of a complete transformer inspection and tracking system that includes a database, preprinted inspection forms and written reports. It runs on The Knowledge Network Computer located in Hartford Steam Boiler's home office and is accessed by our insureds using personal computers and modems.

This paper will discuss the TOGA expert system and its evolution from a prototype system to a comprehensive transformer testing environment.

TRANSFORMER ANALYSIS

Large oil-cooled transformers contain a variety of organic materials such as cellulose solid insulation and mineral oil insulating fluid. These materials deteriorate under the electrical and thermal stresses which exist to some degree in all operating transformers. When oil or cellulose breaks down, certain combustible gases form and dissolve in the oil. The rate and amount of gas generation is important. Normal aging produces gasses at a slow rate; however, incipient or newly forming faults generate gasses at an accelerated rate. These faults also have characteristic energy loads and therefore yield different gas profiles. The dissolved gasses can be identified and quantified using gas chromatography.

A transformer failure expert can review the results of gas chromatography and identify faults occurring in a transformer.

WHAT TOGA IS

TOGA is a knowledge based computer system that emulates the reasoning of a human expert in the analysis of chromatography data to detect faults in oil-cooled transformers. It consists of more than 250 rules that our transformer expert developed during a career analyzing the relationships between dissolved gas concentrations and incipient faults.

TOGA provides the expert with a preliminary analysis and recommendation about the transformer. The expert then looks at additional factors, such as the transformer's age or history, to make a final decision about the condition of the transformer. In this way, TOGA screens good transformers from bad ones, and allows the expert to focus on those transformers needing more immediate attention. Thus, the TOGA system does not replace the transformer expert, rather it enhances his/her productivity.

Paralleling the methods of the human expert, TOGA looks for gas concentrations above and between particular threshold values, and at the relative concentrations of some of these gases. Based upon these "observations" the program determines the nature and severity of the fault, and recommends action to be taken and an appropriate resampling period.

TOGA also analyzes screen test results. It looks at the dielectric strength, the power factors at ambient and elevated temperatures, the acidity, and the interfacial tension, and evaluates the condition of the oil. If necessary, TOGA will recommend the type of preventive maintenance that should be performed. It may recommend that the transformer be resampled before taking action. For instance, if the power factors indicate free water in the sample, there may have been water in the sample bottle.

THE EVOLUTION OF TOGA: THE EXPERT SYSTEM IS EVALUATED

Preventing losses is important to Hartford Steam Boiler and our customers. Therefore, much of our effort and our premium dollars are directed toward developing and maintaining loss prevention programs.

In 1984 Hartford Steam Boiler performed an extensive evaluation of our transformer testing program to determine if it was cost effective. The evaluation identified a threshold transformer size of 5,000 KVA or larger where significant benefits could be accrued. A rigorous analysis was performed in which experienced claims adjusters estimated the cost of the potential loss associated with each discovered fault.

The study estimated an averted loss benefit to Hartford Steam Boiler of $3.00 for each $1.00 spent. Additional benefits would accrue to our customers for amounts below their deductibles.

The cost savings indicated that the program should be expanded to include more transformers. A review of statistics related to transformer oil samples showed that although 75% of those transformers being tested exhibited no problems, every oil analysis report had to be personally reviewed by our transformer expert. This time-consuming process constrained expansion of the program. We had two options: we could either add more transformer experts to our staff, or find ways to increase the productivity of our current expert.

About the same time, Hartford Steam Boiler was becoming more involved in artificial intelligence. We were considering ways the technology might be used to enhance our loss prevention programs. We considered an expert system to assist our transformer expert in the routine screening of oil tests.

The application appeared promising. It met all of the critical criteria needed for a successful implementation of expert system technology. These criteria are discussed in depth in the paper titled "INTERVIEW, A Program to Evaluate Expert System Applications." (1)

The problem domain was well-bounded -- analyzing oil samples to monitor the condition of a transformer. The specific problem task -- identifying incipient faults -- had clearly identifiable inputs (gas concentration data) and output (arcing, corona, etc.) and was well-defined.

There was an adequate source of expertise. Our expert was available and he was willing to participate in the project.

The application was potentially cost effective. If successful, an expert system's assistance in separating those transformers with faults from those without faults could eliminate the need for the expert to review 75% of the test reports. Thus, he would be able to review three times as many transformers as he could without the aid of this expert system.

The project had management's support. The long-term benefits of knowledge preservation and increased productivity were weighed against the short-term impact on our expert's productivity. Management felt that a person of our current expert's caliber could not be found easily. He would need to train new experts in order to expand our transformer testing capacity. Thus his productivity would be adversely impacted in either case.

Management saw the benefit of expert systems and felt we needed to learn how to develop them. It decided the transformer oil testing program was a good place to start. Full management support was given and the Transformer Oil Gas Analyst expert system project was begun.

THE EVOLUTION OF TOGA: THE SYSTEM IS DEVELOPED

TOGA was developed using RuleMaster™. RuleMaster is a software tool kit created by Radian Corporation, a subsidiary of Hartford Steam Boiler, for the development and delivery of expert systems. A key feature of RuleMaster is its ability to build rules from examples. Each example has an unique set of input conditions and an associated outcome. RuleMaster analyzes these input conditions and outcomes and induces "if-then-else" rules which describe the logic captured in the examples.

Rule Induction

In order to understand rule induction, let's look at the process of rating restaurants. Assume that restaurants are rated on the basis of two criteria -- price and atmosphere. Given examples of restaurants, some rated bad, some rated good, and some rated excellent; one can induce or infer the rules used to rate them. These rules associate criteria values (atmosphere and price) with ratings (bad, good, and excellent.) Once the rules are known, they can be used to rate other restaurants according to price and atmosphere.

The following (simplistic) examples are given:

1. Quick-Carrots has a poor atmosphere and low prices, it is a bad restaurant.
2. Quaint-Cakes has a good atmosphere and low prices, it is a good restaurant.
3. Quiet-Candles has a good atmosphere and high prices, it is an excellent restaurant.
4. Quirky-Croissants has a poor atmosphere and high prices, it is a bad restaurant.

From these examples, the following rules about rating restaurants can be induced:

1. If it has a poor atmosphere, it is a bad restaurant.
2. If it has a good atmosphere and low prices it is a good restaurant.
3. If it has a good atmosphere and high prices it is an excellent restaurant.

These rules can now be used to rate any restaurant based on its price and atmosphere.

The next step would be to gather examples and induce rules for the criteria themselves. For instance, what are the criteria for judging atmosphere? (Noise and lighting might be used.) What are some examples of restaurants having a good atmosphere? (Quiet-Candles is quiet and the lighting is soft, it has a good atmosphere.) What rules determining atmosphere can be induced from the examples? (If the noise is quiet and the lighting is soft then the atmosphere is good.)

Developing TOGA's Rules

The first step in building the TOGA system was to identify the possible causes for transformer failure that can be detected by dissolved gas analysis. A knowledge engineer worked with the expert to identify the following types of incipient transformer faults: corona, arcing, thermal overheating due to overloading, and thermal overheating due to either contact resistance or circulating currents in the core of the transformer.

Further discussions identified the criteria the expert was using to detect each of these different faults. For instance, the concentration of acetylene is an indicator of arcing.

Once the faults and criteria were identified, the expert gave examples of actual oil test analyses. The examples associated criteria values with detected faults. The knowledge engineer used RuleMaster to induce from these examples the rules the expert uses for analyzing oil tests. These rules map the relationships between gas concentration profiles and incipient transformer faults.

To illustrate this, the set of examples in Figure 1 shows how a simple rule for corona detection might be constructed. The rule determines whether a corona is unlikely, possible, or likely. The decision is based on four criteria: the concentration of hydrogen, the presence of thermally generated gases, the ratio of hydrogen to acetylene, and the estimated temperature at which the hydrocarbon gases were generated.

The concentration of dissolved hydrogen gas ("H2") may be high, medium, or low, according to ranges set by the expert. (Note: these ranges are dependent on the biases introduced by the sampling methods, extraction methods, and equipment calibration. They may differ from one laboratory to another.) Thermally generated hydrocarbon gases ("THERMAL") may be absent, slight, or present. The hydrogen to acetylene ratio ("COR_RATIO") may be above or below 4. The temperature at which hydrocarbon gases were generated ("TEMP") may be low, moderate, or high.

A hierarchy of rules is supplied by the expert to determine the value of each of these attributes, which fundamentally depend on the dissolved gas concentrations. A "-" value for any attribute indicates that the example is valid for all possible values of that attribute. For instance, the first example in Figure 1 states that a corona is possible when the hydrogen level is high, the ratio of hydrogen to acetylene is above 4, and the temperature is moderate, for all levels of thermally generated gases.

The diagnostic rules induced from the examples in Figure 1 are shown in Figure 2.

A fundamental understanding of the process is:

1. Identify a 'result'. For instance, a TOGA result is an incipient fault such as corona.
2. Identify the criteria that indicate such a 'result'. For instance, the concentration of hydrogen is one indication of corona.
3. Induce rules from examples of criteria values and associated results. For instance, oil tests and their associated faults, as diagnosed by the expert, were used as examples in the TOGA system.

This process was recursively applied to determine gas value thresholds, incipient faults, and locations. The method was then applied to develop the screen test portion of the program.

TOGA was then tested with real data. It was put to work analyzing all of the oil samples being taken. The transformer expert continued to analyze each of these samples. The results of the expert system were compared with the expert's analysis. These validation tests showed that TOGA's identification of faulty transformers agreed with that of the expert 99% of the time. Furthermore, actual problem diagnosis agreed with the expert more than 90% of the time.

THE EVOLUTION OF TOGA: THE DATABASE IS EVALUATED

As we developed TOGA, it became apparent that much of the expert's analysis was based not only on the static values of the gas for a given transformer, but also on trends in the gas values from one test to another. Thus, each time he reviewed the results of a transformer test, he would have to search his paper files to find the reports on the previous tests for that transformer. This was a tedious process and particularly difficult when previous sampling dates and identification numbers were left out of the reports. A database that interfaced with TOGA would provide the expert with easy access to the historical trending data he needed.

One problem we were having with our transformer program was inconsistencies in transformer data. Each time a transformer is tested, transformer nameplate data is written on the sample form by the field representative. This nameplate data is then entered into the computer. This process left much room for human error, transposed numbers, illegible handwriting, or inconsistent spelling. For instance, GE, G.E., and General Electric - can all be interpreted to mean the same manufacturer by anyone familiar with the acronym. However, a computer has difficulty recognizing that these three all refer to the same manufacturer.

A database would greatly enhance the transformer program by providing a source of consistent transformer information to both the human expert and the expert system. It could be used to "pre-print" the sample forms, so that all of the transformer nameplate and policy information would appear on the form. In addition to greatly increasing the data integrity, it was estimated that this would save the field representative between 7 and 20 minutes per transformer. The analyst at the lab would no longer have to key repetitive information, a savings of about 5 minutes per test.

In addition, a database would enhance the entire transformer testing process in a number of other ways. It could be used to schedule and track testing. It could also be used for analyses of different transformer trends, such as correlations among increasing gas concentrations and transformer age. A database provides easy data manipulation to sort and examine data in almost any manner of interest, such as typical gas values, or differing values based on manufacturer.

Thus, as the transformer testing program grew, the benefits of a transformer database motivated the design of the TOGA database.

THE EVOLUTION OF TOGA: THE DATABASE IS DEVELOPED

Before designing the database, we studied the information flow of the transformer program and considered the many functions the database would serve. With this global perspective, we designed the transformer database to be highly flexible, able to meet a wide variety of informational needs.

The TOGA database was implemented with a relational database management system. A relational database organizes information in tables and allows easy access and retrieval of data on an ad hoc basis. The database stores all of the information relevant to the TOGA system: gas chromatography data, screen test data, and transformer nameplate information. In addition, it holds company, policy, address, contact, invoicing, and account information. It also keeps track of other transformer related activity, such as electrical testing. Thus, the database serves a wide audience. Account team members, inspectors, supervisors, engineers, and others, as well as the expert system, can use the database for their specific informational needs.

The TOGA database is designed to optimize data consistency. Maintaining the integrity of a database becomes an increasingly difficult problem as the volume of data grows and when there is a large number of people manipulating the data. For instance, if the same transformer is stored in two different tables and the serial number is changed, it is necessary to ensure that the change occur in both tables. Relational databases can be modeled to avoid storing data redundantly. In addition, "integrity checks" or rules for data entry can be policed by the system. The TOGA database includes a number of these integrity checks. For example, a transformer must have an acceptable policy number associated with it. A policy number is acceptable if it already exists in the policy table. The design also makes use of special validation tables. These tables are, in effect, lists of legal values. For instance, TOGA has a valid manufacturer table. This table stores all the valid spellings of manufacturers that will be accepted by the database. This table contains General Electric but not G.E. These integrity checks and validation tables maintain meaningful and consistent data in the database, and ensure accuracy and completeness when performing data manipulations.

The database provides a number of query and report options. A query is a question that is asked of a database. It retrieves information from the database in a useful format. The expert system uses queries to obtain the test data it needs when making an analysis.

TOGA users also use queries to retrieve information from the database. For example, "What were the gas data values for the last four tests of transformer X?" "What tests were performed between dates X and Y for policy number Z?" "How many screen tests were performed this month?" Thus, TOGA users do not need to be database experts to extract data from the database. They simply choose a query and provide values for the variables. For instance, in the first query above, the user would give a specific serial number for the variable "X."

The database generates printed reports from the user's queries on request. These reports are used for invoicing and work management as well as for data analysis.

The database also assists in the generation of letters to customers. These letters are composed by the expert after compiling information from a number of paper files. Database reports now make this task easier by providing a single source of data. In the future, some of these letters will be composed automatically by the expert system using rules about composing letters and information obtained from the database.

The database has become an important part of TOGA. The expert system interfaces directly with the database, extracting oil and screen test data and storing the results of its analysis. In the future, it will obtain historical and nameplate data from the database and apply new rules associated with trend analysis and transformer age. The transformer expert uses the database for trend analysis and letter writing. Account engineers, and field representatives use the database to monitor the service we are providing our customers. Lab analysts use the database for invoicing.

Thus, the incorporation of a database into TOGA enhances the expert system and increases the efficiency of the transformer testing program.

THE EVOLUTION OF TOGA: INTEGRATION WITH THE TRANSFORMER TESTING PROGRAM

The transformer testing process begins when a Hartford Steam Boiler field engineer draws a sample from an oil-cooled transformer. The sample, together with a form containing customer and transformer specific information, is then sent to Radian Analytical Services (RAS) located in Austin, Texas.

At RAS, laboratory technicians perform the necessary gas chromatography and screen tests. Using a personal computer and a telecommunications software package, they dial into the Hartford Steam Boiler Knowledge Network Computer (KNC) in Hartford, Connecticut and enter the site information and test results into the TOGA database.

At this point, the TOGA expert system is applied to the new data. The results of the analysis are displayed within seconds and are also stored in the database. For those analyses requiring immediate attention, the transformer expert is automatically notified. An electronic message is sent to the expert in Hartford, notifying him that the analysis has been completed.

The expert uses the database to evaluate the transformer's condition by looking at the expert system results, transformer nameplate data, and the results of previous samples. He notes and analyzes any dangerous trends in the gas concentration data and generates a report to the customer.

The expert system recommends a period for resampling the transformer based on its analysis. This recommendation is stored in the database and used to schedule sampling. Those transformers found to be normal are automatically recommended for resampling in one year. If there are indications of incipient faults, the system will recommend more frequent resampling. The expert can override the expert system's recommendation if he does not concur.

Periodically a report is sent to each of our field offices indicating which transformers are due for resampling. Soon, sample forms will be also be generated by the TOGA system. These forms will be preprinted with transformer nameplate information and sent to the inspector upon demand.

THE EVOLUTION OF TOGA: CUSTOMER ACCESS

Many of Hartford Steam Boiler's insureds perform their own transformer testing but either do not have an expert on-site or their expert is overburdened with analyses. Several of our customer's asked us if they could use TOGA because the same benefits that TOGA brought us could apply to them.

It is known that gas chromatography results can differ from one laboratory to another for the same oil sample. Although different laboratories may generate different results for the same sample, results are usually standardized within a laboratory. Therefore the reasoning behind the analyses will not differ, but the threshold values will. For instance, in one laboratory a C2H2 level of 35 ppm may be considered high, while in another, a level of 5 ppm would be high. In both cases however, a high level of C2H2 is an indicator of arcing.

The TOGA system was 'calibrated' to be used with the RAS Laboratory. This means that the threshold values for the gases are consistent with results from this laboratory. Any laboratory equipment that generates data values consistent with those obtained at RAS can be used with the TOGA program. However, results that are inconsistent with the RAS laboratory equipment may be misinterpreted by the TOGA expert system.

A future enhancement to the system could enable laboratory specific calibration of the threshold values. Until then, we caution all users of TOGA of the potential for mistaken analysis, with any gas values obtained in laboratories inconsistent with RAS.

TOGA is just one of the expert systems available through The Hartford Steam Boiler's Knowledge Network Computer. The Knowledge Network Computer is a collection of software and hardware that resides in Hartford Steam Boiler's home office.

The Knowledge Network Computer contains knowledge of machinery trouble shooters, transformer experts and other Hartford Steam Boiler specialties. Authorized users accesses this network by using a personal computer or a terminal and a modem to 'dial-in' to the network via the telephone. We provide all the necessary software, even a program that will perform the set up and dial the telephone. Simple menus guide users to access TOGA or other expert systems. The user also has access to electronic mail.

The Knowledge Network Computer's electronic mail facility gives users the opportunity to communicate directly with Hartford Steam Boiler's experts. If they have any questions about TOGA or concerns about an analysis they can "mail" a message directly to our expert. Our expert can also respond to their questions via the electronic mail.

You can read more about the Knowledge Network Computer in the paper titled: "TURBOMAC: Network Delivery of Problem Solving Knowledge."(2)

FUTURE DIRECTIONS

TOGA, like most expert systems, will never be complete. Now that the basic knowledge of the system has been implemented, the next step is to provide additional functionality for the system's users and audience. We are currently enhancing the database with more reporting features and developing the preprinted forms.

In the future, the expert system will acquire knowledge from the expert about how trending is used, and how to consider additional factors such as the age and manufacturer of the transformer. With the integration of the database, as a source of historical data, rules can now be added to make note of dangerous trends in gas concentrations and to know manufacturer specific problems.

Additionally, the expert system will be expanded to work with the database to perform automatic reporting functions. For instance, it will be used to generate summary reports for the expert. It will also be enhanced to write intelligent letters using data stored in the database. In these letters the expert system would group transformers together by company and draw appropriate attention to those transformers with indications of faults.

The evolution of TOGA has given us a good look at the many potential uses and benefits of an expert system. We have learned that an expert system works well as part of as an evolutionary step in an existing process. In this case, TOGA, facilitated the expansion of Hartford Steam Boiler's existing transformer testing program. The expert system, however, is only one aspect of a complete human and computer environment. While it may improve the consistency and productivity of a human expert it will never learn as much or reason as completely about problems as the expert himself. We have learned that an expert system, when well-designed to assist some known process, is not the end to meet all means, but the means to many ends.

H2	THERMAL	COR_RATIO	TEMP		
high	-	above_4	moderate	=>	corona is possible
medium	absent	above_4	moderate	=>	corona is possible
high	-	above_4	high	=>	corona is unlikely
medium	absent	above_4	high	=>	corona is unlikely
high	-	above_4	low	=>	corona is likely
medium	absent	above_4	low	=>	corona is likely
medium	present	-	moderate	=>	corona is unlikely
medium	slight	-	moderate	=>	corona is unlikely
low	-	-	-	=>	corona is unlikely
-	-	below_4	-	=>	corona is unlikely

Figure 1.
Expert Example For Corona Detection

```
IF the cor_ratio IS "above_4":
    IF temperature IS "low":
        IF level of  H2 IS "low":
            THEN corona is "unlikely"
        IF the level of H2 IS "medium" OR "high"
            THEN corona is "likely"
    IF temperature IS "moderate":
        IF level of H2 IS "low":
            THEN corona is "unlikely"
        IF the level of H2 IS "medium":
            IF thermally generated gases ARE "absent"
               THEN corona is "possible"
            IF thermally generated gases ARE "slight" OR "present"
               THEN corona is "unlikely"
           ELSE corona is "possible"
       ELSE  corona is "unlikely"
   ELSE  corona is "unlikely"
```

Figure 2.
Rules Induced From Examples Shown In Figure 1

REFERENCES

1. Howes, J.R., INTERVIEW, A Program to Evaluate Expert System Applications, Expert Systems Applications for the Electric Power Industry, June 5-8, Orlando, Florida, Sponsored by Electric Power Research Institute. (To be published.)

2. Klimczak, B., TURBOMAC: Network Delivery of Problem Solving Knowledge, Expert Systems Applications for the Electric Power Industry, June 5-8, Orlando, Florida, Sponsored by Electric Power Research Institute. (To be published.)

GESTAL: A Specialized Tool to Build Real-Time Alarm Processing and Fault Diagnosis Expert Systems for Power Network Control Centers

J.-M. ARÈS and P. GIROUARD
Ages Intelligence Ltd.
P.O. Box 655, Place du Parc,
Montreal, Quebec, Canada, H2W 2P2

ABSTRACT

In recent years, the task of power system operators has become more complex as a result of the large amount of information generated by modern Energy Management Systems (EMS). In many instances, the overwhelming amount of information presented during network disturbances results in a longer operator response time. In order to alleviate this problem, Ages Intelligence has developed GESTAL™, a specialized tool to build and maintain real-time expert systems for alarm processing and fault diagnosis in power network control centers. A prototype of GESTAL and an associated expert system were developed and validated using Lisp and ART™. A more elaborate version of the tool has been implemented in a C/OPS83® environment. A pilot expert system for twelve substations is currently ongoing both off-line and on-line testing at Hydro-Québec.

1. INTRODUCTION

Following a disturbance in a power network, control center operators must analyze sequences of alarm messages in order to establish a fault diagnosis. Based on this diagnosis, the operators can take the necessary actions to ensure network stability and/or to restore the load. In instances where the number of alarm messages is considerable, the operators face a complex analysis problem which may be time consuming. Such a delay can be costly to the utility since the load is not restored immediately and since

ART™ is a trademark of Inference Corporation.
GESTAL™ is a trademark of Ages Intelligence Ltd.
OPS83® is a registered trademark of Production System Technologies, Inc.

certain types of faults may propagate if appropriate actions are not undertaken in time. On the other hand, the operators cannot precipitate their actions and perform manœuvres based on a superficial analysis of the alarm messages since a false manœuvre may, in certain instances, result in considerable equipment damage or in the propagation of the fault. Therefore, considering the substantial amount of information which may be generated by modern Energy Management Systems during crisis situations, the need for real-time fault diagnostic systems becomes eminent.

The problems of alarm processing and fault diagnosis in power network control centers, along with related expert system prototypes, have been presented in [1, 2, 3, 4, 5]. Most of these papers discuss expert system techniques to perform alarm processing/fault diagnosis without proposing a solution for the large-scale implementation of such expert systems. Furthermore, none of these papers propose a solution which takes into consideration the temporal nature of the problem. This paper presents GESTAL, a tool to deploy real-time expert systems that integrate alarm processing and fault diagnosis capabilities. The tool incorporates reasoning strategies to overcome the problems of temporal reasoning and of performance degradation resulting from the large number of alarm points being monitored. Furthermore, the development and maintenance of the knowledge bases are greatly simplified by a specialized knowledge base compiler.

2. DESIGN OBJECTIVES

The functional and system specifications of GESTAL were elaborated by two knowledge engineers through discussions with control center operators and power network design engineers. The main design objectives which were identified are presented below:

a) Simple interpretation of the generated diagnoses:

The fault diagnoses should present only the information which is essential to assist the operator identify the root-cause and the consequences of the fault. In addition, detailed explanations of the obtained diagnoses should be available upon request.

b) Automatic analysis:

The expert systems should be designed such that no user interaction is required to obtain analysis results; all of the needed parameters should be obtained directly from the EMS data base. This feature is highly desirable as the operator should not be burdened with an additional task in crisis situations.

c) Real-time performance:

The fault diagnoses should be generated fast enough to allow the operator to take corrective actions. A single expert system should be able to monitor in the order of 100 000 alarm points. Consideration should be given to the fact that in crisis situations, Energy Management Systems are capable of generating over 500 alarms per minute [6].

d) Robustness of diagnostic capabilities:

The inference strategy should be able to cope with the fact that status messages may not be available for every relay in the network, and that, during disturbances, certain status messages may not be received due to data acquisition problems. Furthermore, if the received data justify more than one interpretation, the expert system should present the various possibilities.

e) Flexibility of the knowledge base:

The expert system should be capable of supporting the analysis of alarms from substations of different configurations. Furthermore, it should be able to diagnose the operation of the various types of relay protection and recovery systems that exist in the network.

f) Simple maintenance procedures:

A standard methodology should be specified to allow non-computer experts to maintain the knowledge base. Moreover, the architecture of the expert system should support gradual up-scaling.

3. ARCHITECTURE

Based on the design objectives, the model-based architecture illustrated in figure 1 was developed. The GESTAL tool consists of four basic components: the Analysis Module, the Programming Interface, the User Interface, and the Communication Interface. A GESTAL expert system is built with the Programming Interface by defining a frame-based model for each substation from which alarms are to be analyzed. Essentially, the substation models contain knowledge describing the characteristics and the behavior of the relay protection and recovery systems. The central component of the expert system is the Analysis Module.

It contains the inference engine, the rules and the procedural code that define the alarm processing and fault diagnosis strategies. The Communication Interface is used to obtain the relevant information from the EMS data base whereas the User Interface presents the analysis results in an ergonomic menu-driven environment.

Analysis strategy:

One of the major challenges in developing an automatic diagnostic feature is to devise a reasoning strategy which can define the proper time interval for the analysis of any given alarm sequence. Since alarm sequences correspond to the signature of physical events whose duration may vary, it is crucial to be able to identify when sufficient information has been received to generate a diagnosis. Figure 2 illustrates this problem: the set of messages $s_i = \{a_3, a_4, a_5\}$ may correspond to the signature of either event e_1, e_2, e_3, or e_4. Hence, if the set of alarms s_i corresponds to event e_4, the reasoning mechanisms must recognize this and consider alarms a_1 through a_9 before generating a diagnosis. In order to overcome this problem, the reasoning strategies utilized by GESTAL expert systems dynamically specify the time window for the analysis according to the alarm messages that are received. Basically, as illustrated in figure 3, this *Dynamic Time Windowing* technique is implemented as follows: as alarm messages are received, the analysis module gradually constructs directed graphs in which a node represents an alarm message and an arc represents a causal or an associative relation. Obsolete alarm messages and inconsistent diagnostic graphs are discarded whereas accepted and completed diagnostic graphs are translated into natural language format and presented to the operator.

In order to ensure that the real-time performance remains independant of the number of alarm points being monitored, the Analysis Module's inference strategies also incorporate a focus of attention method that dynamically controls which portions of the knowledge base are invoked based on the messages received. This data-driven approach is extremely important considering that a single expert system must be able to monitor in the order of 100 000 alarm points.

Maintenance:

Considerable attention was given to the issues of maintenance and expansion of the knowledge base. In order to ensure the robustness of the fault diagnosis systems throughout their life cycle, a knowledge representation strategy in which the expert systems can be expanded and/or updated without altering the procedural knowledge base (Analysis Module) was adopted. A simple structured language was defined

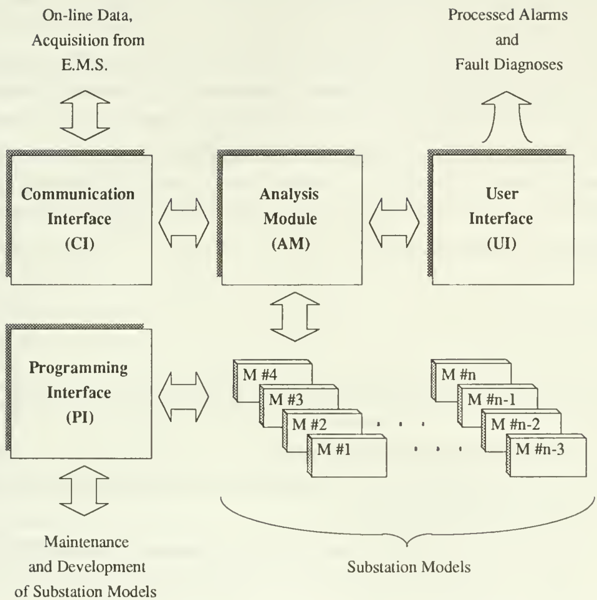


Figure 1: Achitecture of a GESTAL expert system.

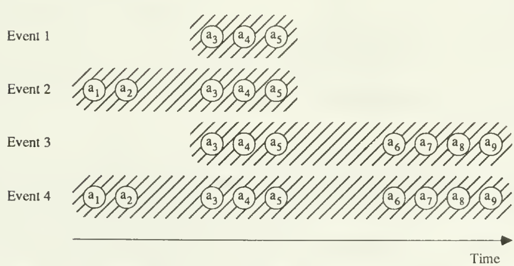


Figure 2: Definition of the proper time window for the analysis.

to model the required substation and network specific knowledge. Accordingly, modifications which reflect changes in substation or network configuration simply involve editing and compiling a portion of the declarative knowledge base (Substation Models) through the Programming Interface. The modular configuration of the declarative knowledge base along with the static nature of the procedural knowledge base ensure that the integrity of the overall system is preserved even in the presence of minor discrepancies in the Substation Models. The knowledge incorporated into these models can be easily extracted from the alarm point descriptions and from the schematics describing the protection and recovery systems. Furthermore, very little computer background is required to be able to modify the knowledge base. In brief, a fault diagnosis expert system can be developed incrementally and the acquisition of knowledge can be done according to a standard methodology.

4. EXAMPLE

The primary role of GESTAL based expert systems is to help power system operators assess correctly and more rapidly the cause(s) and the consequence(s) of network disturbances in order to reduce the delay required to take proper corrective actions. However, the format of the generated fault diagnoses and explanations is also well suited for use in the contexts of post-fault analysis and operator training. The fault diagnoses contain the following information:

- *Fault identification:* the type of fault and the affected component(s) are identified. Depending on the resolution of the received information, either the exact fault stimulus or a set of possible stimuli is presented.
- *Relationships between multiple faults:* when appropriate, the expert system establishes relationships between faults that are currently being diagnosed and one or more previously diagnosed fault(s).
- *Description of system operation:* the expert system describes the exact sequence in which protection and recovery systems have operated.
- *Resulting state:* when appropriate, the expert system presents the resulting state of affected components.

Each fault diagnosis is justified by a set of alarm messages and these explanations can be displayed to the operator upon request. The GESTAL tool also incorporates some traditional alarm processing features such as alarm prioritization and identification of false alarms through algorithmic methods. The following example illustrates some of the functional characteristics of GESTAL based systems.

Consider figure 4, illustrating a portion of a power network, and suppose that in substation A, a differential fault activates the primary protection of transformer T1 and that breaker 120-3 is defective. The result is that:

a) Breakers 300-1, 300-2, 300-3, 120-1, and 120-4 trip.

b) Since 120-3 does not trip, the backup protection of T1 is activated and thus breakers 120-2 in substation A and 120-6 in substation B trip to isolate L4.

c) A recovery system in substation B causes breaker 120-4 to close automatically in order to feed T3 and T4 through L3.

A subset of the alarm sequence corresponding to this fault, as well as the fault diagnosis and the explanation generated by the GESTAL expert system are illustrated in figures 5, 6 and 7 respectively. Note that the level of abstraction of the fault diagnosis is such that the operator can rapidly identify the cause and the consequences of the fault. In contrast, the explanation provides a more detailed perspective on how the expert system arrived at each of its conclusions. The justifying evidence is based on the alarm messages received during the disturbance and on the state of certain status points in the EMS data base.

Off-line tests based on data from previous network disturbances have confirmed the accuracy of the reasoning strategies and demonstrated that the response time of GESTAL based systems will be extremely short even in crisis situations involving rates of over 500 alarms per minute. For instance, on a VAXstation II/GPX™, the response time to generate a fault diagnosis has typically been less than one second.

5. SUMMARY AND FUTURE WORK

We have introduced GESTAL, a specialized tool to build and maintain real-time alarm processing and fault diagnosis expert systems for power network control centers. In order to support modular development and simple maintenance procedures of the expert systems, the knowledge required to perform the analysis has been separated into an Analysis Module (procedural knowledge base) and into a set of Substation Models (declarative knowledge base). Moreover, a Dynamic Time Windowing Technique was devised to overcome the problems of temporal reasoning in this expert system application. Test results have demonstrated the accuracy and efficiency of the inference strategies. It is anticipated that these will permit the deployment of large-scale expert systems to monitor in the order of 100 000 alarm points without significant degradation in run-time performance.

VAXstation™ is a trademark of Digital Equipement Corporation.

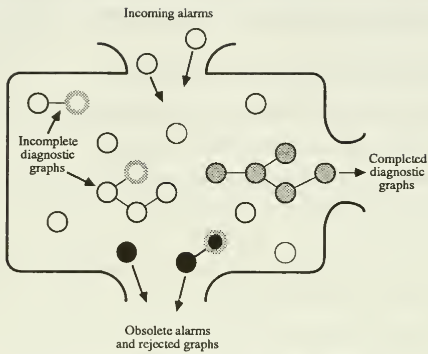


Figure 3: Progressive generation of diagnoses using Dynamic Time Windowing.

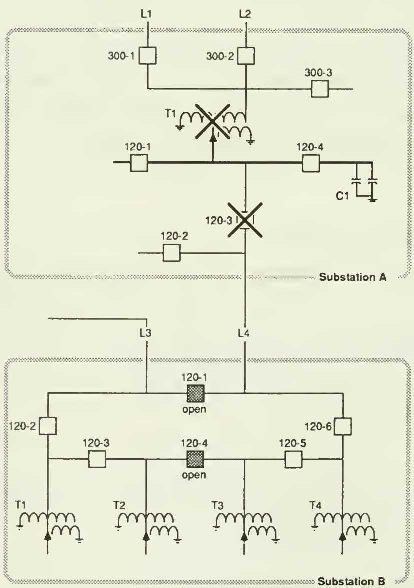


Figure 4: Portion of a power network.

Date	Time	Station	Message	State	Date	Time	Station	Message	State
890415	161507	A	B300-1	-Open-	890415	161507	A	L4-B94	-Alarm-
890415	161507	A	B300-2	-Open-	890415	161507	B	B120-6	-Open-
890415	161507	A	B300-3	-Open-	890415	161507	B	L4-A94	-Alarm-
890415	161507	A	B120-1	-Open-	890415	161507	B	L4-B94	-Alarm-
890415	161507	A	B120-2	-Open-	890415	161508	A	L5--85	-Alarm-
890415	161507	A	B120-4	-Open-	890415	161508	A	L5--85	-Normal-
890415	161507	A	T1--87	-Alarm-	890415	161508	A	T1--87	-Normal-
890415	161507	A	S1--27	-Alarm-	890415	161508	A	T1-94B	-Normal-
890415	161507	A	S1--27	-Normal-	890415	161508	A	L6--85	-Alarm-
890415	161507	A	L5--85	-Alarm-	890415	161508	A	L6--85	-Normal-
890415	161507	A	L6--85	-Alarm-	890415	161508	A	L4-A94	-Normal-
890415	161507	A	L5--85	-Normal-	890415	161508	A	L4-B94	-Normal-
890415	161507	A	L6--85	-Normal-	890415	161508	B	B120-4	-Close-
890415	161507	A	CP--49	-Alarm-	890415	161508	B	T3-RS3	-Alarm-
890415	161507	A	CP--49	-Normal-	890415	161508	B	L4-A94	-Normal-
890415	161507	A	T1-94B	-Alarm-	890415	161508	B	L4-B94	-Normal-
890415	161507	A	L4-A94	-Alarm-	890415	162027	B	T3-RS3	-Normal-

Figure 5: Sequence of alarm messages.

890415 161507 Substation A T1.

Fault:

The protection system of T1 in substation A has operated due to:

Differential

Resulting State:

Substation A: T1 off-line.
Substation B: T3 on-line.
Substation B: T4 on-line.
L4: off-line.

Diagnosis:

Substation A: protection of T1 operated abnormally:

Substation A: breaker 120-3 did not trip;

Substation A: backup protection of T1 was activated;

Substation A: protection of L4 operated normally.

Substation B: protection of L4 operated normally.

Substation B: recovery system of T3 and T4 operated.

Figure 6: Fault diagnosis produced by the expert system.

890415 161507 Substation A T1.

Fault:

The protection system of T1 in substation A has operated due to:

Differential (<?> Substation A: T1--87)

Resulting State:

Substation A: T1 off-line (<?> Substation A: T1---V is 0).
Substation B: T3 on-line (<?> Substation B: T3---V is 122).
Substation B: T4 on-line (<?> Substation B: T4---V is 121).
L4: off-line (<?> Substation A: L4---V is 0).
L4: off-line (<?> Substation B: L4---V is 0).

Explanation:

Substation A: protection of T1 operated abnormally:
<?> 890415 161507 Substation A: B300-1 tripped.
<?> 890415 161507 Substation A: B300-2 tripped.
<?> 890415 161507 Substation A: B300-3 tripped.
<?> 890415 161507 Substation A: B120-1 tripped.
<?> 890415 161507 Substation A: B120-3 did not trip.
<?> 890415 161507 Substation A: B120-4 tripped.
<?> 890415 161507 Substation A: T1--87 was received.

Substation A: breaker 120-3 did not trip;

Substation A: backup protection of T1 was activated;
<?> 890415 161507 Substation A: T1-94B was received.

Substation A: protection of L4 operated normally.
<?> 890415 161507 Substation A: B120-2 tripped.
<?> 890415 161507 Substation A: B120-3 did not trip.
<?> 890415 161507 Substation A: L4-A94 was received.
<?> 890415 161507 Substation A: L4-B94 was received.

Substation B: protection of L4 operated normally.
<?> 890415 161507 Substation B: B120-1 was already open.
<?> 890415 161507 Substation B: B120-6 tripped.
<?> 890415 161507 Substation B: L4-A94 was received.
<?> 890415 161507 Substation B: L4-B94 was received.

Substation B: recovery system of T3 and T4 operated.
<?> 890415 161508 Substation B: B120-4 reclosed.
<?> 890415 161508 Substation B: T3-RS3 was received.

Figure 7: Explanation corresponding to the fault diagnosis.

Having successfully addressed the fundamental implementation issues of real-time performance, automatic reasoning, and maintenance of knowledge bases we envisage that the next generation of GESTAL fault diagnosis tools will be integrated either as a built-in feature of an EMS software system or as a standalone microcomputer-based package.

REFERENCES

1. D. S. Kirschen, et al., "Artificial Intelligence Applications in an Energy Management System Environment.", Proceedings of the Symposium on Expert Systems Applications to Power Systems, Stockholm-Helsinki, August 22-26, 1988.

2. S. Talukdar, and E. Cardozo, "A Distributed Expert System for Fault Diagnosis.", Proceedings of IEEE Power Industry Computer Conference, Montreal, Canada, May 1987, pp. 101-106.

3. C. Fukui, J. Kawakami, "An Expert System for Fault Section Estimation Using Information from Protective Relays and Circuit Breakers", IEEE Winter Power Meeting, 1986, Paper No. 86 WM 112-7.

4. B.F. Wollenberg, "Feasibility Study for an Energy Management System Intelligent Alarm Processor ", IEEE Transactions on Power Systems, Vol. PWRS-1, No. 2, May 1986, pp. 241-247.

5. "Artificial Intelligence Technologies for Power System Operations", EPRI Report EL-4323, Electric Power Research Institute, Palo Alto, CA.

6. H. Amelink, A. Forte, R.P. Guberman, "Dispatcher Alarm and Message Processing", IEEE Trans. on Power Systems, Vol. PWRS-1, No. 3, pp. 188-194, August 1986.

[8]

ISBN 1-56032-102-4